MDPI

Review

A Review of Computer Vision-Based Structural Deformation Monitoring in Field Environments

Yizhou Zhuang [1], Weimin Chen [1], Tao Jin [2,3,*], Bin Chen [2,4], He Zhang [2,3] and Wen Zhang [1]

1   College of Civil Engineering, Zhejiang University of Technology, Hangzhou 310014, China; yizhouzhuang@zjut.edu.cn (Y.Z.); weiminchen@zjut.edu.cn (W.C.); wenzhang@zjut.edu.cn (W.Z.)
2   School of Engineering, Zhejiang University City College, Hangzhou 310015, China; chenbin@zucc.edu.cn (B.C.); zjuzhanghe@zju.edu.cn (H.Z.)
3   Department of Civil Engineering, Zhejiang University, Hangzhou 310058, China
4   Yangtze Delta Institute of Urban Infrastructure, Hangzhou 310005, China
*   Correspondence: jintao@zucc.edu.cn

Abstract: Computer vision-based structural deformation monitoring techniques were studied in a large number of applications in the field of structural health monitoring (SHM). Numerous laboratory tests and short-term field applications contributed to the formation of the basic framework of computer vision deformation monitoring systems towards developing long-term stable monitoring in field environments. The major contribution of this paper was to analyze the influence mechanism of the measuring accuracy of computer vision deformation monitoring systems from two perspectives, the physical impact, and target tracking algorithm impact, and provide the existing solutions. Physical impact included the hardware impact and the environmental impact, while the target tracking algorithm impact included image preprocessing, measurement efficiency and accuracy. The applicability and limitations of computer vision monitoring algorithms were summarized.

Keywords: computer vision; structural deformation monitoring; field environment; environmental impact; target tracking algorithm impact

1. Introduction

Transportation infrastructure systems such as bridges, tunnels and railroads are important component systems for national social production and national development. With the tremendous development of social productivity, these transportation infrastructures are tested in two major ways. On the one hand, the tonnage and number of existing means of transportation may exceed the design load-carrying capacity; on the other hand, civil engineering structures including bridges, are subjected to various external loads or disasters (such as fire and earthquakes) during their service life, which in turn reduces the service life of the structures. By carrying out inspection, monitoring, evaluation, and maintenance of these structures, we can ensure the long life and safe service of national infrastructure and transportation arteries, which is of great strategic importance to support the sustainable development of the national economy.

In the past two decades, structural health monitoring (SHM) has emerged with the fundamental purpose of collecting the dynamic response of structures using sensors and then reporting the results to evaluate the structures' performance. Their wide deployment in realistic engineering structures is limited by the requirement of cumbersome and expensive installation and maintenance of sensor networks and data acquisition systems [1–3]. At present, the sensors used for SHM are mainly divided into contact type (linear variable differential transformers (LVDT), optical fiber sensors [4–9], accelerometers [10,11], strain gauges, etc.) and non-contact types (such as global positioning systems (GPS) [12–14], laser bibrometers [15], Total Station [16], interferometric radar systems [17], and level computer vision-based sensors). Amongst the existing non-contact sensors, the GPS sensor is easy to

install, but the measurement accuracy is limited, usually between 5 mm and 10 mm, and the sampling frequency is limited (i.e., less than 20 Hz) [18–22]. Xu et al. [23] made a statistical analysis of the data collected using accelerometers and pointed out that the introduction of maximum likelihood estimation in the process of fusion of GPS displacement data and the corresponding acceleration data can improve the accuracy of displacement readings. The accuracy of the laser vibrometer is usually very good, ranging from 0.1 mm to 0.2 mm, but the equipment is expensive and its range is usually less than 30 m [24]. Remote measurements can be performed with better than 0.2 mm accuracy using a total station or level, but the dynamic response of the structure cannot be collected [25,26].

With the development of computer technology, optical sensors and image processing algorithms, computer vision has been gradually applied in various fields of civil engineering. High-performance cameras are used to collect field images, then various algorithms are used to perform image analysis on a computer to obtain information such as strain, displacement, and inclination. After further processing of these data, the dynamic characteristics such as mode shape, frequency, acceleration and damping ratio can be obtained. Some researchers extract the influence line [27,28] and the influence surface [29] of a bridge structure from the spatial and temporal distribution information of vehicle loads, which are used as indicators to evaluate the safety performance of the structure. However, the long-term application of computer vision in the field is limited in many ways; for example, the selection of targets, measurement efficiency and accuracy, environment impact (especially the impact of temperature and illuminate changes).

This paper reviews the computer vision-based studies in field environments, including the system composition, target tracking algorithms, environmental influencing factors and current achievements. It is organized as follows: Section 2 briefly describes the basic hardware composition of the computer vision monitoring system; Section 3 introduces the flow of monitoring, camera calibration methods, feature extraction and different target tracking algorithms. Section 4 reviews the current application scenarios of computer vision in the field of SHM, analyzes the shortcomings in field applications and lists the corresponding solutions, and finally briefly describes some basic requirements of long-term monitoring. Section 5 makes a summary and points out the important problems that need to be further studied in long-term field applications of deformation monitoring systems based on computer vision.

2. System Composition

A computer vision-based structural deformation monitoring system includes an image acquisition system and an image processing system. The image acquisition system includes a camera, lens, and target to collect video images, while the image processing system performs camera calibration, feature extraction, target tracking, and deformation calculation, which purpose is to process the acquired image and calculate the structural deformation. This section will briefly introduce the basic components of the image acquisition system in the computer vision monitoring system. The image processing system will be introduced in Section 3.

2.1. Camera

The camera is an important part of the image acquisition system, and its most essential function is to transform received light into an electrical signal through a photosensitive chip and transmit it to the computer. Photosensitive chips can be divided into CCD and CMOS according to the different ways of digital signal transmission. The main differences between them are that CCD has advantages over CMOS in imaging quality, but its cost is much higher than that of CCD, so it is suitable for high-quality image acquisition; CMOS is highly integrated and saves electricity compared with CCD, but the interference of light, electricity and magnetism is serious and its anti-noise ability is weak, so it is more suitable for high-frequency vibration acquisition [30,31].

The selection of the camera needs to consider the following points: (1) the appropriate chip type and size is to be selected according to the measurement accuracy and application scenario; (2) because of the limited bandwidth, the frame rate and resolution of the camera are contradictory, so the frame rate and resolution should be balanced in camera selection; (3) industrial cameras appear to be the only option for long-term on-site monitoring.

2.2. Lens

The lens plays an important role in a computer vision system, and its function is similar to that of the lens in a human eye. It gathers light and directs it ontp the camera sensor to achieve photoelectric conversion. Lenses are divided into fixed-focus lenses [32,33] and zoom lenses [34]. Fixed-focus lenses are generally used in laboratories, and high-power zoom lens are generally used for long-distance monitoring such as of long-span bridge structures and high-rise buildings. The depth of field is related to the focal length of the lens; the longer the focal length, the shallower the depth of field.

The following points need to be considered in the selection of shots: (1) a low distortion lens can improve the calibration efficiency; (2) an appropriate focal length for the camera sensor size, camera resolution and measuring distance should be selected; (3) a high-power zoom lens is appropriate for medium and long-distance shooting.

2.3. Target

The selection of targets directly affects the measurement accuracy, and an appropriate target can be selected according to the required accuracy. There are mainly two kinds of target: artificial targets and natural targets. Ye et al. [35] introduced six types of artificial targets [19,36,37] (flat panels with regular or irregular patterns, artificial light sources, irregular artificial speckles, regular boundaries of artificial speckle bands, and laser spots) and a class of natural targets [38,39]. Artificial targets can provide high accuracy and are robust to changes in the external environment, just as artificial light sources can improve the robustness of targets in light and the possibility of monitoring at night. The disadvantage of artificial target is that they need to be installed manually, which may change the dynamic characteristics of the structure. Natural targets rely on the surface texture or geometric shape of the structure, which is sensitive to changes of the external environment, and their accuracy is not high.

The following points should be noted in the selection of targets: (1) when the target installation conditions permit, priority should be given to selecting artificial targets to obtain stable measurement results; (2) the selection of targets should correspond to the target tracking algorithm in order to achieve better monitoring results.

3. Basic Process

The flowchart of deformation monitoring based on computer vision is shown in Figure 1, and can be summarized as follows: (1) assemble the camera and lens to aim at artificial or natural targets, and then acquire images; (2) calibrate the camera; (3) extract features or templates from the first frame of the image, then track these features again in other image frames; (4) calculate the deformation. The following is a brief description of camera calibration, feature extraction, target tracking and deformation calculation.
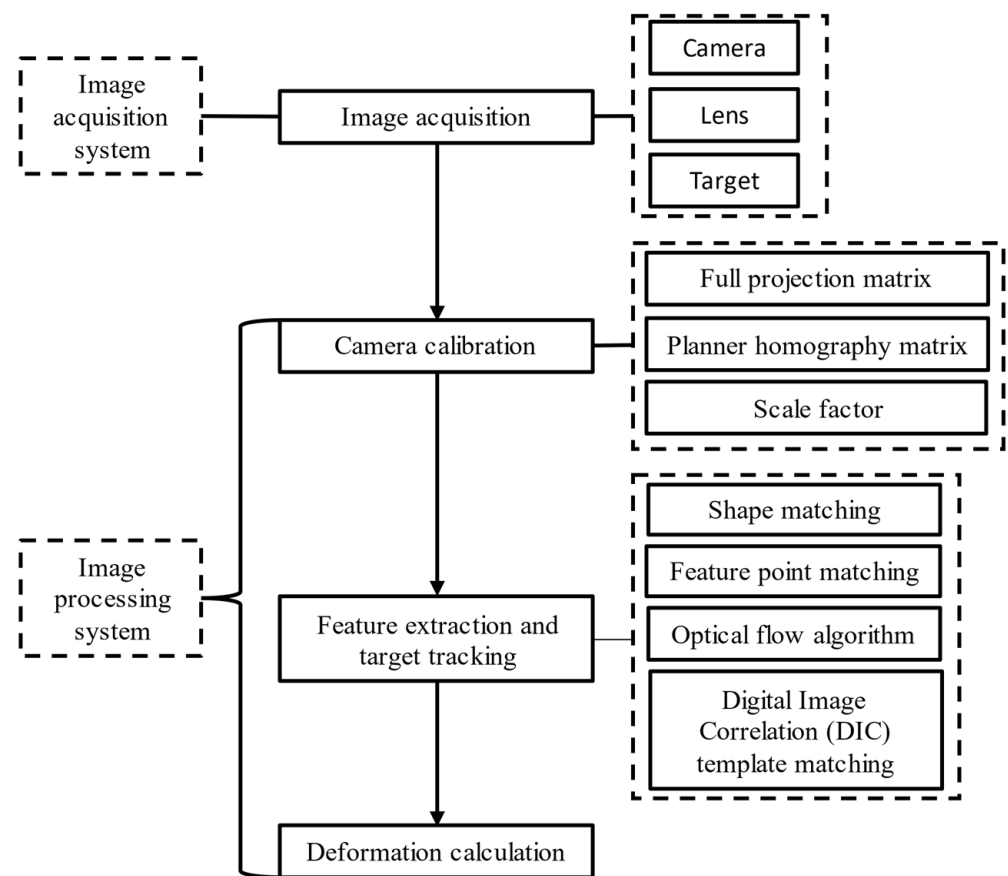
Figure 1. Process of deformation monitoring based on computer vision.

3.1. Image Acquisition

Image acquisition includes these steps: (1) determine the position to be monitored; (2) arrange artificial targets or use natural targets on the measurement points; (3) select an appropriate camera and lens; (4) assemble the camera lens and set it firmly on a relatively stationary object; (5) aim at the target and acquire images.

3.2. Camera Calibration

Camera calibration [40] is the process of determining a set of camera parameters which associate real points with points in the image. Camera parameters can be divided into internal parameters and external parameters: internal parameters define the geometric and optical characteristics of the camera, while external parameters describe the rotation and translation of the image coordinate system relative to a predefined global coordinate system [41]. In order to obtain the structural displacement from the captured video image, it is necessary to establish the transformation relationship from physical coordinates to pixel coordinates. The common coordinate conversion methods are full projection matrix, planar homography matrix, and scale factor.

3.2.1. Full Projection Matrix

The full projection matrix transformation reflects the whole projection transformation process from 3D object to 2D image plane. The camera internal matrix and external matrix can be obtained by observing a calibration board, which can be used to eliminate image distortion and has a high accuracy [42]. Commonly used calibration boards include checkerboard [43] and dot lattice [44].

Figure 2a shows the relationship between the camera coordinate system, the image coordinate system and the world coordinate system. A point T (X, Y, Z) in the real 3D world appears at the position t (x, y) in the image coordinate system after the projection

transformation (where the origin of the coordinates is P). The relationship between the pixel coordinate system and the image coordinate system is shown in Figure 2b. Therefore, the equation for converting a point from a coordinate in the 3D world coordinate system to a coordinate in the pixel coordinate system is

$$
S \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} == \begin{bmatrix} fx & \gamma & u_x & 0 \\ 0 & fy & u_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = M_1 M_2 X \tag{1}
$$

where S is the scale factor from Equation (3), f_x and f_y are the camera lateral axis and vertical axis focal lengths, γ is the angle factor of the lens, u_x and u_y are lateral and vertical offsets of the principal axs, respectively, R is the rotation matrix of size 3×3 and t is the translation matrix of size 3×1, M_1 is the camera internal parameter, and M_2 is the camera external parameter.
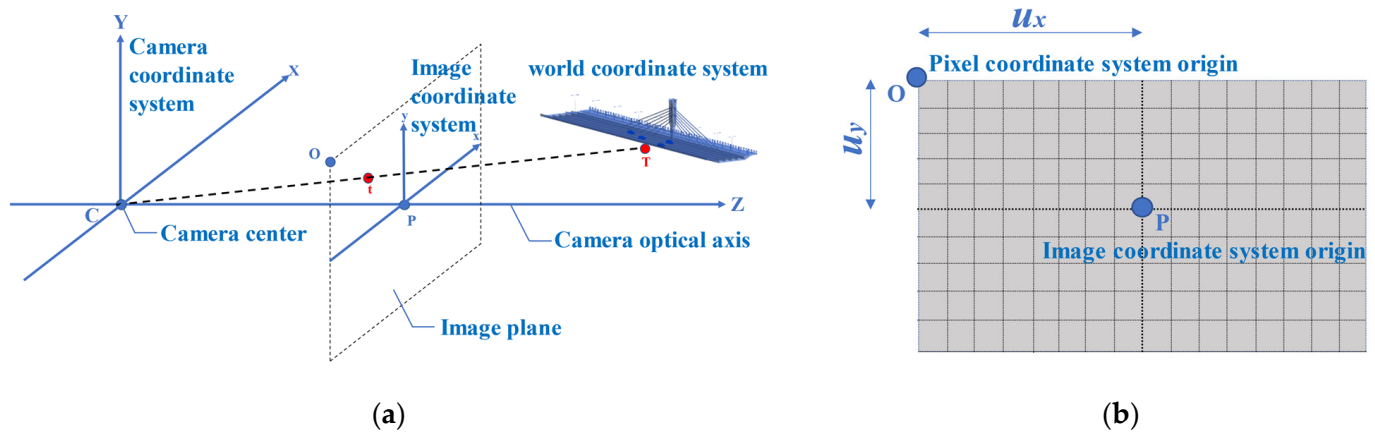

Figure 2. (**a**) Relationship among the camera coordinate system, the image coordinate system, and the world coordinate system; and (**b**) Relationship between the pixel coordinate system and the image coordinate system.

Park et al. [45] and Chang et al. [41] calibrated with T-bar and checkerboard respectively to eliminate the measurement error caused by camera distortion and accurately measure the 3D dynamic response of a structure.

3.2.2. Planar Homography Matrix

In practical engineering applications, the above calibration process is relatively complex. To simplify the process, Equation (1) can be expressed

$$
S \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = K \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{2}
$$

where K is called planar homography matrix [46], which can reflect the relationship between the corresponding points on two images and is not affected by the angle between the optical axis and the structural plane [43].

The planar homography matrix is suitable for the case where there is an angle between the image plane and the moving plane of the object, and the angle is not easy to measure [42]. The position of at least four known points on the moving plane can be used to solve the planar homography matrix. Khuc et al. [29] and Xu et al. [47,48] both used known structural dimensions to solve for the planar homography matrix, construct the corresponding relationship between image coordinates and 3D world coordinates, and estimate the time history information of lateral and vertical displacement of a bridge.

3.2.3. Scale Factor

The scale factor (S) provides a simple and practical calibration method. As shown in Figure 3a, when the camera optical axis is perpendicular to the surface of the object, S (unit: mm/pixel) can be obtained based on the internal parameters of the camera (focal length, pixel size) and the external parameters of the camera and the surface of the object (measurement distance) in a simplified calculationfrom the simplified formula

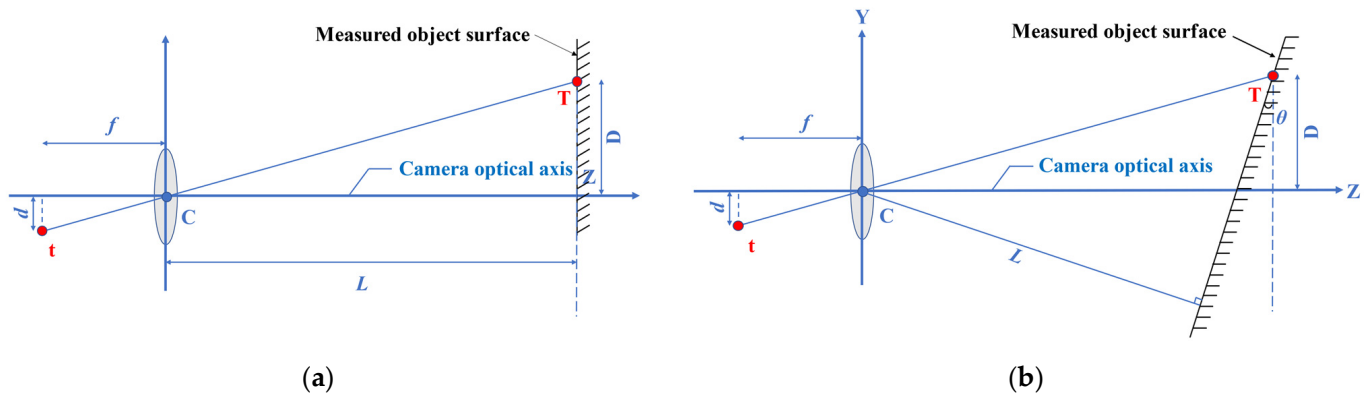$$S = \frac{L}{f}d_{pixel} = \frac{D}{d}d_{pixel} \tag{3}$$

(**a**)　　　　　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 3. Scaling factor calibration method. (**a**) Camera optical axis orthogonal to measured object surface; (**b**) Camera optical axis intersecting obliquely with measured object surface.

When the optical axis of the camera is not perpendicular to the surface of the measured object (as shown in Figure 3b), the included angle would affect the measurement accuracy [49]. Feng et al. [1] studied the influence of different angles between the optical axis and the surface of the measured object on the accuracy, and found that S can be determined by:

$$S = \frac{L}{f \, \cos^2 \theta}d_{pixel} = \frac{D}{d \, \cos^2 \theta}d_{pixel} \tag{4}$$

where f represents the focal length; L represents the distance from the camera to the measured object surface along the optical axis, also known as object distance; D represents the distance from the measuring point to the optical axis; and d represents the distance from the measuring point on the image to the origin.

References [50–55] build S according to known physical dimensions on the surface of the object (such as the dimension of an artificial object or the dimension of the structural member obtained from the design drawing) and the corresponding image dimensions to measure the displacement of the structure.

Among these camera calibration algorithms, the appropriate coordinate conversion method needs to be selected according to the field environment and measurement purpose. The full projection matrix and the planar homography matrix have no restraint on camera position but need a calibration plate. The full projection matrix is suitable for 3D deformation monitoring, and the planar single response matrix and scale factor are suitable for 2D deformation monitoring.

3.3. Feature Extraction and Target Tracking

Feature extraction is used to obtain the unique information in the image (such as shape features, feature points, grayscale features, and particle features). The purpose of target tracking algorithms is to find these features again in other image frames. Common target tracking algorithms in civil engineering structural deformation monitoring include shape matching, feature point matching, optical flow estimation and digital image correlation (DIC) template matching.

3.3.1. Shape Matching

In an image, shape is a description of an edge or region, and shape matching is an image matching algorithm to identify and locate measured objects through image edge features. There are many algorithms for edge detection, such as Zernike operator [56], Roberts operator, Sobel operator [57], Log operator [58], Canny operator [59] and generalized Hough algorithm [60]. Among them, the Canny operator is widely used because of its high performance [61,62].

The principle of shape matching is relatively simple and can be used for displacement monitoring of structures with obvious shapes. The advantages are: (1) the calculation is relatively simple and the matching speed is fast; (2) it is robust to change of illumination because it tracks the geometric boundary of the object; (3) this measurement has an advantage for linear structures such as slings.

3.3.2. Feature Point Matching

Feature point matching is a target tracking method based on feature extraction and matching. The key points in computer vision are those which are stable, unique and invariant to image transformation, such as building corners, connection bolts, or other shaped targets [63,64]. The common methods of feature point detection include Harris Corner [65], Shi–Tomasi Corner [66], scale invariant feature transform (SIFT) [32,67], speed-up robust feature (SURF) [68], binary robust independent elementary features (BRIEF) [69], binary robust invariant scalable keypoint (BRISK) [70], and fast retina keypoint (FREAK) [71].

A feature point matching algorithm needs to select appropriate feature descriptors according to the measurement object to describe feature points mathematically and carry out image registration. It is usually suitable for structures with rich textures or certain shapes (Such as circle, hexagon or rectangle). Feature point matching has the following characteristics: (1) it deals with the whole image area and has accurate matching performance [72]; (2) it extracts texture features of the structure and is not sensitive to illumination and shape transformation; (3) the greater the number of feature points used, the higher is the precision (however, this increases the calculation time).

3.3.3. Optical Flow Algorithm

Optical flow algorithm is an image registration technique in which the surface motion in a three-dimensional environment is approximated as a two-dimensional field by using the spatio-temporal pattern of image intensity [73]. The optical flow algorithm can accurately provide the velocity and displacement of the object by tracking the trajectories of pixels, but it has great limitations and makes the following assumptions [74]: (1) the brightness of objects in adjacent frames remains unchanged; (2) the motion of objects in adjacent frames is small enough; (3) the motion between adjacent pixels is consistent [75]. Common optical flow algorithms include Lucas–Kanade [76,77], Horn–Schunck method [78], Farneback method [79], block match method [80], and phase-based optical flow [45,81,82]. Among those, Lucas–Kanade is fast and easy to implement, and it can perform motion tracking in the selected measurement area, especially of robust feature points, while other algorithms need to calculate every pixel in the image, which is slow.

The optical flow algorithm is similar to the feature point matching algorithm in that it tracks feature points on the image and prefers target patterns with distinct and robust features over the whole test period. The optical flow algorithm has the following characteristics: (1) target features need to be clear; (2) sensitivity to illumination changes; (3) only motion components perpendicular to local edge direction can be detected, such as bridge cable vibration; (4) optical flow describes the motion information of the image brightness and is more suitable for measuring dynamic displacement.

3.3.4. DIC Template Matching

The basic principle of DIC is to compare the same points (or pixels) recorded between two images before and after deformation, and to calculate the motion of each point [83]. As

a representative non-interference optical technique, DIC has the advantage of continuous measurement of the whole displacement field and strain field. It is a powerful and flexible surface deformation measurement tool in experiments on solids, and it has been widely accepted and used [84–87]. If we track only a small pixel area, we can track and monitor the displacement of the measuring points of the structure [88,89], which is called template matching. The basic process of monitoring displacement by template matching is as follows [90–92]: (1) select some areas of the first frame image as templates; (2) use these templates to scan line by line in a new image frame; (3) then use the relevant criteria to match the degree of similarity and determine the pixel coordinates of the matched template; (4) calculate the pixel displacement and convert it to the actual displacement.

The relevant criteria include the following six mathematical algorithms: (1) cross-correlation (CC); (2) normalized cross-correlation (NCC); (3) zero-normalized cross-correlation (ZNCC); (4) sum of squared differences (SSD); (5) normalized sum of squared differences (NSSD); and (6) zero-normalized sum of squared differences (ZNSSD) [93].

In computer vision-based displacement measurement, the NCC matching method is the most popular, and there are numerous applications of the method. Template matching based on DIC has the following characteristics: (1) it is not very robust to light changes, slight occlusions, and scale changes; (2) an artificial target is beneficial to improve the success rate of matching; (3) huge computational expense during the template matching; calculation in the frequency domain can save computation time.

3.4. Deformation Calculation

Deformation computation is the process of transforming pixel displacement into actual displacement. First, high quality images are collected; then, 3D motion in the real world is decomposed into planar motion by camera calibration; later, the matching algorithm is used to track and calculate the pixel distance of the target moving in the image plane. Finally, the pixel distance is converted into proportional actual distance.

The accuracy of displacement depends not only on the camera calibration method and target tracking algorithm, but also on the environment, so the influence of environment on the accuracy of displacement calculation needs to be understood. This is the problem that needs to be solved in current field applications. The most important thing is to improve the algorithm so that it can adapt to the changing environment.

4. Computer Vision-Based Deformation Monitoring in Field Environment

Computer vision-based sensors have made great strides in the lab, and computer vision-based monitoring systems have the following advantages over conventional attached sensors and other non-contact optical sensors: (1) providing displacement measurements in both time and frequency domains [94]; (2) measuring multiple targets simultaneously [95]; (3) non-contact long-distance high-precision measurement; (4) simple setup and lower labor intensity [96].

Although computer vision-based structural deformation monitoring has broad prospects, there are still some challenges and problems to be studied. Up to now, there have been few cases in which the structural deformation monitoring system based on computer vision could be used stably in structural health monitoring for a long time. In the process of indoor experimental application research, a controllable experimental environment enables the image acquisition system to stably collect high-quality images or video files and to get better results by post-processing. However, in on-site long-term monitoring, the erection conditions of targets and cameras are limited by the on-site environment through factors such as target installation difficulty, environmental vibration, measuring distance, and image acquisition and transmission rate. Image processing also needs to meet the requirements of long-term stable real-time monitoring under uneven changes of temperature and light, occlusion, and real-time processing of image data, and to output reliable displacements. These challenges and problems will be important parts to be considered in future research and engineering practice.

4.1. Application Scenario of Computer Vision Monitoring System

Researchers began to use video analysis in civil engineering structure monitoring and achieved results. At present, computer vision-based monitoring technology has been applied to many fields of SHM, including bolt looseness detection [97–99]; quantifying disaster impact [100–104]; cable force monitoring [55,105–108]; modal frequency monitoring [2,19,52,63,74,95]; structural deformation measurement in 2D [34,109–112] and 3D [50,113]; bridge structural influence line [27,28] and influence surface measurement [29]; structural damage detection and localization [51,52]; updating finite element models of structures [114,115]; static and dynamic rotational angle measurement of large civil structures with high resolution [116], and spatio-temporal distribution of traffic loads [117].

In practice, researchers have found the influence of computer vision-based monitoring system on practical applications. According to the system composition, these influencing factors are divided into two aspects: physical influence research and target tracking algorithm influence research. Physical influence corresponds to the image acquisition system, and target tracking algorithm corresponds to the image processing system. Sections 4.2 and 4.3 will focus on the impacts of these two aspects, analyze their influence mechanisms, and present some solutions.

4.2. Hardware Impact and Environmental Impact

4.2.1. Hardware Impact

The camera and targets are the main parts of the image acquisition system. In field applications, we need to consider the stability of the long-term use of the camera and solve the problem that target cannot be installed on some structures.

- Camera

Shutter mode and photosensitive chip size can lead to a difference in imaging. The rolling shutter method may cause image distortion when recording fast-moving objects. When using this type of camera, the effect of this distortion on the measurement results should be corrected, and a global shutter can solve this problem [110,118]. The bigger the photosensitive chip, and the higher the picture pixel density, the higher the theoretical measurement accuracy, but the higher the economic cost. In addition, camera heating will cause chip heating, resulting in errors. Ma et al. [119] conducted an in-depth study on the strain measurement errors caused by self-heating of CCD and CMOS cameras. When the temperature increases, the virtual image expansion will cause a 70–230 $\mu\varepsilon$ strain error in the DIC measurement, which is large enough to be noticed in most DIC experiments and hence should be eliminated.

The inherent frequency of a structure determines the sampling frequency of the camera. According to Nyquist's Theorem, when the sampling frequency is less than twice that of the measured signal, aliasing (i.e., false low-frequency components in the sampled data) may occur. Different sampling frequencies should be adopted for different structures: a rigid structure needs high sampling frequency, while a flexible structure can permit a lower sampling frequency. This not only can save energy, but also can reduce calculation and allow real-time monitoring. For example, if the highest frequency at which significant (visibly detectable) motion of the bridge structure occurs is below 10 Hz, then a sampling rate of 20 Hz should be sufficient to avoid aliasing.

- Target

The target is key to the accuracy of computer vision measurement. There are two types of targets: artificial targets and natural targets. The artificial target is an obstacle in the field application of current computer vision-based measurement methods. It must be attached to the surface of the measured object. The installation of artificial targets may require equipment such as a bridge inspection vehicle, which is not only time-consuming, but also unsafe. In addition, the installation of artificial targets may change the dynamic characteristics of the structure [74]. Brownjohn et al. [112] studied the effect of the properties

of targets (including both artificial targets and edge features of the structure). The results show that the noise of the vision sensor is inversely proportional to the size of the target.

Ehrhart et al. [120] attached a circular target to a pedestrian bridge to measure the bridge vibration, and proved that, for a single frame structure and an observation distance within 30 m, a motion larger than 0.2 mm can be detected. Khuc et al. [32,64] proposed a new vision-based displacement measurement method that did not require installation of manual markers and instead used robust features extracted from the image as virtual targets. Fukuda et al. [2] and Ye et al. [121] used feature matching between continuous images to realize displacement measurement. Yoon et al. [38] introduced a target-free approach for vision-based structural system identification using the Kanade–Lucas–Tomasi (KLT) tracking algorithm and Shi–Tomasi corners. This work could accommodate multi-point displacement measurement of a six-story building model in the laboratory; however, it did not provide verification with conventional displacement sensors. Dong et al. [74] extracted virtual markers from images using robust feature detection algorithms that represent texture or other unique surface features of the structure, and can select the best markers according to different scenarios. They thus made the matching algorithm more adaptive, and verified the effectiveness of the algorithm by measuring structural vibrations of soccer stadium bleachers. Kim et al. [105] carried out environmental vibration tests on the Gwangan Bridge in South Korea to measure the sling structure motion without any target to verify the effectiveness of the non-target strategy in the measurement of the dynamic characteristics of bridge hanger cables.

At present, feature point matching and shape matching are mainly used in target-free strategies, because these two methods can effectively utilize original features of the structure and have robustness to illumination variation.

4.2.2. Environmental Impact

When it comes to field monitoring applications, a variety of external environmental factors that are rare in the laboratory, such as temperature change, camera movement caused by environmental vibration, illumination change, and illumination mutation caused by shielding, will lead to the increase of image noise and the decrease of matching accuracy. In order to solve these problems, researchers have made efforts to reduce the systematic errors caused by these factors. This section first classifies these environmental influences according to their error mechanisms and then summarizes the current solutions.

- Optical refraction

Refraction is a complex optical phenomenon occurring naturally, and vision sensors are easily affected by optical refraction at high temperature, so it is still a challenge to monitor a structure remotely using vision sensors. It can be observed that the change of air density caused by heating causes a change of optical refractive index, which leads to distortion of video images and thus error in displacement measurement. Optical refraction error caused by uneven heating of the air is shown in Figure 4. During a field test, when the air between the camera and the target structure was heated unevenly, measurement error increased as the measurement distance increased [1].
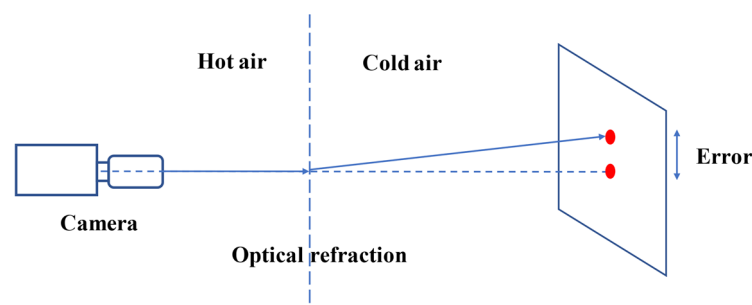

Figure 4. Principle of optical refraction error.

At present, the research on optical refraction mainly focuses on techniques related to static image restoration (that is, processing image distortion). There are many ways to reduce distortion in a single image, such as lucky image, using region-level fusion based on the dual tree complex wavelet transform [122], multi-frame super-resolution reconstruction [123], B-spline-based nonrigid registration [124], and derivative compressed sensing. However, these techniques are custom-made for static images and therefore do not apply to distinguishing structural motion. Luo et al. [125] used a normal random distribution to fit the error caused by optical refraction, which reduces the displacement measurement error caused by optical refraction by about 67.5%. Luo et al. [126] comprehensively studied the characteristics of distortion and displacement error caused by hot air, established a hot air error model, and quantified the measurement error caused by hot air through bridge displacement measurement experiments carried out in high-temperature weather.

Up to now, the influence of optical refraction on vision-based measurement has rarely been mentioned. However, the error caused by optical refraction can reach 50 mm, which shows the importance and necessity of this research. According to Owens's [127] research, the refractive index of air varies with air pressure, air temperature and air composition. Therefore, when considering the influence of thermal haze, the effects of humidity and air pressure should also be considered.

- Camera motion

In addition to temperature change, the position of the camera will change due to various influencing factors (e.g., traffic load, thermal expansion and cold shrinkage of brackets, or loose structures). When measuring a real structure outdoors, the position and direction of the camera often change slightly due to wind, vibration and ground instability [128]. Ye et al. [50] believed that the thermal expansion and cold contraction of the mounting bracket would also cause a small change in the position of the camera, and this small motion of the camera, even a very small rotation, would lead to a very large error with increasing range. Figure 5 is a vision-based monitoring system in which the target is fixed, and Figure 5a,b show the errors caused by camera translation and rotation, respectively. It can be seen that, when the camera translates a distance d_y along the vertical optical axis, the error is d_y; However, when the camera rotates slightly through an angle d_θ, the error caused is $L \cdot \tan d_\theta$, which is unacceptable. Therefore, in order to improve the accuracy of absolute displacement estimation, several camera motion subtraction techniques have been developed: (1) digital high-pass filtering (DHF) [15,129]; (2) background modification (BM) [51,106,130]; (3) inertial measuring unit (IMU) [131]; and (4) ego-motion compensation (EC) [132].
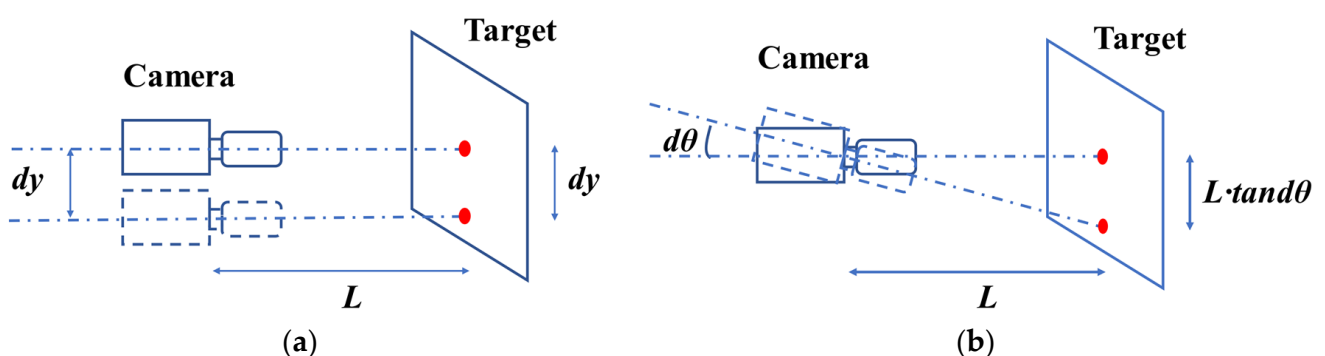

Figure 5. Analysis of error due to camera motion: (**a**) camera translation; and (**b**) camera rotation.

At present, DHF and IMU are mainly used in the field of unmanned aerial vehicle (UAV) displacement measurement, in which DHF can eliminate UAV flight frequencies the by digital high-pass filtering. Garg et al. [15] removed the low-frequency component of the UAS using a high-pass Butterworth filter, and measured the displacement response of a railway bridge under train load; the estimated peak and RMS errors were under 5% and 10%, respectively. An IMU consists of a DC gyroscope and accelerometers that can

respond to very low frequency (almost 0 Hz) vibrations. Ribeiro et al. [131] estimated the displacement and rotation of a UAV by numerical integration, and measured a static concrete structure with a peak value error of 1.47 mm (15.5% relative error) and an RMS error of 9.3%.

BM is a simple and convenient method which uses fixed objects such as buildings and mountains as reference points in the background to calculate the relative displacement of the target [51,106,130]. This method is effective to measure the absolute displacement of a structure during the flight of a UAV. Yoon et al. [118] reproduced the vertical dynamic displacement of a pin-connected steel truss bridge undergoing revenue-service train traffic for 250 s in the laboratory using a servo-hydraulic motion simulator. By tracking background characteristics, the motion of the UAV was modified, and the root mean square error of the corrected displacement was reduced from 116 mm to 2.14 mm. Yoneyama et al. [128] measured the plane rotation and translation of a rigid body and changed the camera position and angle. Even if the camera rotation angle is more than 30°, the error is less than 0.1 mm, which verified the feasibility of deducing multi-degree-of-freedom motion of the camera through background correction. Chen et al. [133] measured the antenna at the top of a building with the building itself as a reference point, and measured the relative motion of the antenna. Compared with the frequency measured by a laser vibrometer at close range, the error was less than 1.7%. Khaloo et al. [134] used a distance-based outlier detection method based on Chebyshev's theorem [135] to accurately estimate the flow vector of pixels in the static background region and subtracted this from the structural pixel flow vector to correct the undesired camera motion.

When the background conditions cannot be satisfied, Lee et al. [132] proposed a long-term displacement measurement strategy that uses a sub-camera to aim at a fixed target near the installation position and calculates the relative motion of the dual-camera system to correct the displacement measured by the main camera. This technical reduced the motion error from 44.1 mm to 1.1 mm.

DHF and IMU may be more suitable for the displacement measurement using drones, because drones always fly with a low-frequency vibration. For long-term monitoring in a relatively stable environment, where there will be no low-frequency vibration, these two techniques have high computational complexity and high cost. Therefore, they are not suitable for long-term monitoring. BM and EC are relatively simple to calculate and have good adaptability to different scenarios, and researchers have carried out field experiments for up to four months to verify their effectiveness [50,132]. These are solutions that can be considered for widespread use.

- Illumination changes and partial occlusion

Optical refraction, camera motion, illumination change and partial occlusion all lead to measurement errors, but the mechanisms are different. Optical refraction and camera motion cause measurement errors, but do not cause changes in image quality, while illumination change and partial occlusion will lead to image quality degradation or even failure to produce a usable image, resulting in the inability to find the target correctly. Deformation monitoring methods based on computer vision are easily affected by environmental conditions when they are applied on a project site [136].

Shi et al. [66] proposed a feature selection criterion based on how the tracker works, which is an optimality criterion based on construction, through which occlusion and point mismatched features can be detected. Yuan et al. [137] proposed a new interpretation standard, which achieves an accuracy of 0.01~0.02 mm under different light intensity conditions and is more accurate than the NCC method. Ullah et al. [138] proposed orientation code matching (OCM), which is based on matching the gradient information around each pixel in the form of an orientation code and is robust to background changes and to illumination fluctuations caused by shadows or highlights. Feng et al. [51,111,114] developed the OCM algorithm into an application program for civil engineering structures. Through indoor shaking table tests, the frequency of a frame structure and the displacement time history of a railway bridge under train load were extracted, and the

robustness of the proposed vision sensor to adverse environmental conditions such as low light, background image interference and partial template occlusion was verified. This is of great significance to the development of computer vision in civil engineering displacement monitoring. Luo et al. [139,140] proposed a new edge enhancement matching technique, which can extract both gradient magnitude and gradient direction at the same time, and can track low-contrast features robustly. This technique was verified on a 16.9 m span steel girder railway bridge and a 448 m span steel suspension bridge. Lee et al. [33] firstly used adaptive region of interest (ROI) cropping to narrow the search range, obtained the marker boundary through an edge detection filter, and verified the robustness of the algorithm to illumination by field experiments. Khuc et al. [64] used a geometric transformation method to discard the outliers in the matching pool, so as to reduce the problem of incorrect matching due to frequent illumination changes and monitor the vibration of a stadium stand structure.

In order to overcome the influence of illumination change and partial occlusion, deep learning methods were introduced. Lichao et al. [141] proposed a convolutional neural network structure to learn the adaptive target template update strategy for a given initial template, cumulative template and current frame template. Xu et al. [47] proposed a new algorithm that integrates depth learning, a convolutional neural network and correlation-based template matching. This algorithm covers adjacent regions by changing the size and local movement of the template region and can adapt to drastic changes of the target pattern. It was verified in short-range and long-range monitoring activities, considering background change, illumination change and shadow effects. Dong et al. [142] used spatio-temporal context learning and Taylor approximation to track the target, and verified the robustness under illumination change and fog interference.

Summarizing the above-mentioned methods, there are currently two main types of methods to solve the light transformation and partial occlusion problems: (1) using image gradient magnitude and direction to extract image edge information for target tracking, which is robust to illumination changes and partial occlusion; and (2) using deep learning methods to train the template and background, which can adapt to constant changes in illumination and background.

4.3. Impact of Target Tracking Algorithm

Target tracking algorithms directly determine the speed and accuracy of displacement calculation, and depend on the acquisition of high-quality images. The images collected in the field are often accompanied by significant noise (such as Gaussian noise, exponential distribution noise, Rayleigh noise, uniform distribution noise, or salt and pepper noise), and this noise will lead to reduced target tracking accuracy. This section introduces (1) field image preprocessing; (2) real-time performance of the algorithm; (3) accuracy of the algorithm; and (4) the choice of algorithm and how to balance accuracy and efficiency.

4.3.1. Image Preprocessing

The quality of the image directly affects the accuracy of the recognition algorithm, so preprocessing is needed before image analysis. The main purpose of image preprocessing is to simplify the data to the maximum extent so as to improve the reliability of feature extraction, image segmentation, matching and recognition. In field applications, external factors such as sudden changes of illumination, rain, and environmental vibration cannot be controlled. In order to estimate the dynamic characteristics of the structure in this case, image processing technology is especially important [105]. The common digital image processing technology in computer vision includes image transformation, image coding compression, image enhancement and restoration, image segmentation and image description.

In order to correct the geometric distortion between a deformed image and an unde-formed image, Kim et al. [14] developed an image processing algorithm which reduces the noise in the frequency domain by natural frequency analysis and accurately measures the dynamic response of a sling. In order to obtain concrete information on steel can-

tilever beam damage, Song et al. [52] first performed noise filtering by using a discrete wavelet transform, and then provided precise damage localization by using a continuous wavelet transform. Javh et al. [143] combined accelerometer and camera data and used the complex frequency domain least square method to avoid the burden of noise data from the high-speed camera and to measure the modal frequency of a scaled cantilever model. Kim et al. [106] applied spatial image enhancement technology, a smoothing filter and a sharpening filter to reduce image noise and also improve the fuzzy part, which improved the recognition rate of a sling, resulting in a measured cable force error less than 1.1%.

At present, image processing technology has been widely used in the field of vision-based monitoring. However, for different application scenarios, it is an unsolved problem to determine the settings of many parameters (filter, binarization, image pyramid). In addition, excessive image processing requires lengthy calculations, which makes application of image preprocessing techniques to vision-based measurement a challenge.

4.3.2. Measurement Efficiency

Most of the short-term laboratory and field applications are post-processing of recorded images or video files to obtain structural deformation information. On the one hand, the algorithm cannot process the collected image files in real time; on the other hand, it can analyze the saved video files many times to get satisfactory results. While a long-term monitoring system requires real-time and stable output of structural deformation information, the computational efficiency of algorithms can limit the application of vision sensors when high-frequency measurements and simultaneous measurements at multiple points need to be obtained.

Lecompte et al. [144] investigated the effect of the size of a subset of the scatter pattern on the measured in-plane displacement efficiency, and showed that the larger the subset, the higher the efficiency. Guizar-sicairos et al. [87] proposed the upsampled cross correlation (UCC), whose registration accuracy is very accurate for nonlinear optimization algorithms, but greatly reduces the calculation time and memory requirements. Zhang et al. [145] proposed an improved Taylor approximation refinement algorithm and a subpixel localization algorithm, both of which are at least five times faster than UCC. Feng et al. [1,114] limited the search area to the predefined ROI near the template position in the previous image, and then processed the ROI region frame by frame using UCC and OCM algorithms. Through shaking table tests of a frame structure, it was verified that a vision sensor could quickly track the multi-point displacement time history of artificial targets or targets on the structure with a maximum RMS error of 0.72%. Dong et al. [74] used a sparse optical flow calculation method (Lucas–Kanade method), which greatly reduces the number of calculated pixels compared with the global optical flow, and accurately measures the dynamic response of the grandstand structure of a football stadium under crowd load. Guo et al. [146] proposed an improved inverse synthesis algorithm based on Lucas–Kanade, which can complete a displacement extraction in 1 millisecond without the need to install any pre-designed targets on the structure.

In addition to increasing the computational efficiency of the algorithm, improving the hardware configuration can also increase the measurement efficiency, but raises the cost.

4.3.3. Measurement Accuracy

Measurement accuracy is an important indicator of deformation monitoring performance. Unlike indoor measurement, on-site measurement often requires long-distance measurement, and the measurement accuracy decreases with increase of the measurement distance. Under certain hardware conditions, the sub-pixel estimation method can solve the measurement accuracy problem to some extent. Pan et al. [147] summarized subpixel methods, including the coarse–fine search method, double Fourier transform, genetic algorithms, artificial neural network methods, correlation coefficient curve-fitting or interpolation, Newton–Raphson iteration, and gradient-based methods, and pointed out that the latter three methods are the most commonly used due to their simplicity, effectiveness and accuracy. The subpixel method can solve the problem of detection accuracy, and it

is necessary to improve the computational efficiency of subpixel accuracy for practical engineering application.

MacVicar-Whelan et al. [148] and Jensen et al. [149] respectively proposed linear interpolation and nonlinear interpolation to improve image resolution, which improves the measurement accuracy to some extent. Bruck et al. [150] proposed a new digital image processing algorithm (Newton–Raphson iteration algorithm), and their experiments show that the Newton–Raphson iteration algorithm can determine displacement and displacement gradient more accurately than a coarse-fine search method. Pan et al. [151] combined an inverse compositional matching strategy with Gauss–Newton without sacrificing subpixel accuracy, and proposed the inverse-compositional Gauss–Newton (IC–GN) algorithm, which is 3 to 5 times faster than the Newton–Raphson iterative algorithm, with an accuracy of less than 0.0222 pixels in the x and y directions. Tian et al. [93] used an efficient and accurate IC–GN algorithm to track a target point, monitored multi-point displacement on a steel truss highway railway bridge, and achieved an accuracy of 0.57 mm at 288 m. Qu et al. [152] proposed an edge detection method combining a pixel level method (Sobel operator) and a subpixel level method (Zernike operator), which is much faster than the Zernike operator, but the detection accuracy is close to that of the Zernike operator. Zhang et al. [145] integrated improved Taylor approximation refinement and localization refinement into vision-based sensors and measured the vibration of a high-speed railway noise barrier. These two improved algorithms (Taylor approximation refinement: RMS error 0.61%, localization refinement: RMS error 0.73%) are at least 5 times faster than traditional UCC (RMS error 0.75%) when the accuracy is similar. Dong et al. [42] combined SIFT feature point and Visual Geometry Group (VGG) descriptor (SIFT–VGG) algorithms as a strategy for vision-based displacement measurement. This integrated strategy improves the measurement accuracy of the original SIFT method by 24% and greatly improves the accuracy of displacement recognition. Fukuda et al. [19] used the azimuth obtained by bilinear interpolation to achieve sub-pixel resolution when measuring structural vibration (vibration frequency: 0.1 Hz~50 Hz, vibration amplitude: 50 mm), with the standard error reduced from 0.14 mm to 0.043 mm. Feng et al. [1] used different levels of subpixel accuracy and pixel accuracy comparison in the laboratory, proved that there is a linear relationship between subpixel accuracy and subpixel level, and accurately measured structural vibration of less than 1 mm. Mas et al. [153] proved the realistic limit of sub-pixel accuracy through a simple numerical model, and found that the maximum resolution enhancement and dynamic range of the image can be achieved.

In these studies, it is found than subpixel can make the measurement accuracy reach a higher level. Many studies have shown that subpixel accuracy varies on the order of 0.5 to 0.01 pixels [147,153]. In theory, subpixel algorithms can achieve the maximum accuracy of displacement calculation, but in practical applications, where real experimental images may be contaminated by many factors such as environmental vibration, temperature change, camera heating, and illumination change, the accuracy often fails to reach ideal levels. In addition, few reported quantitative works have been performed to systematically evaluate their subpixel registration accuracy and computational efficiency or have attempted to solve an existing discrepancy. It is necessary to understand the limitations and performance of these sub-pixel registration algorithms.

4.3.4. Suggestions for Field Application Algorithms

Reviewing the development of application research and engineering practice of SHM based on computer vision, we find it has made great progress. Until now, computer vision has been continuously developed and applied to various fields of civil engineering monitoring. In the application of structural health monitoring based on computer vision, target tracking is the most important step, which directly determines the efficiency, accuracy and reliability of vision sensors. However, the performance of target tracking algorithms depends on application scenarios, and researchers have found applicability and limitations of different algorithms in engineering practice. Table 1 summarizes the advantages and limitations of these algorithms and lists their application scenarios.

Table 1. Characteristics, limitations and application scenarios of the proposed tracking algorithms.

Algorithms	Characteristics and Restraints	Application Scenarios
Shape matching	High efficiency, real-time monitoring, robust properties to non-uniform illumination and partial edge blur, distinct geometry features; short-distance monitoring	Cable structure, tower, long-span bridge, stadium
Feature point matching	High efficiency, high accuracy, robust to illumination change, distinct geometry features; rich surface texture, uncertain number of feature points, short-distance monitoring	Stadium structure, footbridge, railway bridge, urban bridge
Optical flow algorithm	Full field displacement, natural target, suitable for motion tracking; short-distance monitoring, sensitive to illumination change and partial occlusion	Stadium structure, footbridge, railway bridge, urban bridge
DIC template matching	Long-distance and short-distance monitoring; artificial target, sensitive to illumination change	Cable structure, long-span bridge, stadium structure, footbridge, railway bridge, urban bridge, tower

4.4. Measurement Results

Table 2 classifies the problems from Sections 4.1–4.3 and presents the current solutions and the results achieved in practical applications. However, these measurement results were determined in specific environments, and actual measurement results must be determined according to the camera resolution, measurement distance, lighting environment and target contrast used in the field.

Table 2. Measurement results.

Research Point	Reference	Algorithms	Test Description	Results
Target	Ehrhart et al. [120] (2015)	Shape matching (artificial target)	Shaking table test, least squares fit of ellipse, precision quantitative evaluation of accuracy	At the distance of 6 m and 31 m, the error is less than 0.01 mm and 0.2 mm, respectively
	Tian et al. [93] (2016)	DIC template matching (artificial target)	Field test, DIC technology based on IC-GN, displacement~time history curve	At the distance of 300 m, the average error is 0.5674 mm
	Feng et al. [115] (2015)	Template matching (natural target)	Railway bridge test, modal identification, modify finite element model	At the distance of 9 m, the vision sensor and the accelerometer measure the exact same first-order frequency
	Khuc et al. [64] (2017)	Feature point matching (natural target)	Stadium grandstand structure test, feature extraction with Hessian matrix, dynamic displacement measurement	At the distance of 3 m and 13 m, the error is less than 0.01 mm and 0.04 mm, respectively
	Khuc et al. [154] (2020)	Feature point matching (natural target)	Tower test, Canny edge detection and Hough transform, modal identification	At the distance of 1.84 m, first-order natural frequency error below 2%
Camera motion and optical refraction	Garg et al. [15] (2019)	Digital high-pass filtering	Shaking table test, dynamic displacement measurement	At the distance of 4 m, the maximum error is between 10~15%, and the RMS error is between 2~5%
	Ye et al. [50] (2021)	Background modification	Long-term field monitoring, 3D structural deformation measurement	Eliminate the error caused by thermal expansion and cold contraction of camera bracket
	Lee et al. [132] (2020)	Ego-motion compensation	Long-term field monitoring, displacement measurement	Measurement error is reduced from 44.1 mm to 1.1 mm
	Ribeiro et al. [131] (2021)	Inertial measuring unit	Experiment test, modal identification, dynamic displacement measurement	The maximum error of displacement measurement is 1.47 mm and RMS error is 9.3%
	Luo et al. [125] (2020)	Adaptive optical-turbulence error filter	Field test, displacement~time history curve	Measurement errors are significantly reduced by about 67.5% from 0.0845 to 0.0275 mm

Table 2. *Cont.*

Research Point	Reference	Algorithms	Test Description	Results
Illumination change and partial occlusion	Ribeiro et al. [21] (2014)	Artificial light source	Field test, using artificial light source, displacement ~ time history curve	At the distance of 15 m and 25 m, the error is less than 0.1 mm and 0.25 mm, respectively
	Feng et al. [111] (2015)	A matching algorithm based on the gradient information	Field test, a matching algorithm based on the gradient information, dynamic response of steel and concrete bridges under train load	At the distance of 30.48 m, the maximum displacement error is 2.83%, and the average error is 1.39%
	Shan et al. [155] (2015)	Shape matching	Cable force measurement, shape matching, cable modal identification	The first three frequencies of free vibration of stayed-cable model are accurately measured
	Dong et al. [142] (2019)	A matching method based on Spatio-Temporal Context Learning	Experiment test, a matching method based on Spatio-Temporal Context Learning, illumination change and fog interference, displacement~time history curve	The proposed subpixel estimation method is faster than UCC by about 50 times
	Xu et al. [47] (2021)	Combined with deep learning	Field test, displacement~time history curve	Centimeter-level accuracy can be achieved at distances of more than 715 m
Image preprocessing	Kim et al. [105] (2013)	Image transform technology	Ambient vibration tests, suspension bridge hanger cables, dynamic response and modal frequencies	The error of measuring sling modal frequency and cable force is within 0.5%
	Kim et al. [106] (2013)	Image enhancement techniques	Ambient vibration tests, smoothing filter and sharpening filter, stay cables, dynamic response	The error of measuring suspension bridge hanger cables natural frequency and cable force is within 2%
	Tian et al. [62] (2019)	Image description and segmentation technology	Impact test, Hough transform based on gradient, modal parameters identification	The recognition rate of vibration mode is more than 84%

Table 2. *Cont.*

Research Point	Reference	Algorithms	Test Description	Results
Measurement efficiency and accuracy	Qu et al. [152] (2005)	Edge detection method using Sobel-Zernike moments operator	numerical tests	The accuracy reaches 87.75% of the sub-pixel level, and the speed is increased by 5 times
	Pan et al. [156] (2011)	DIC template matching based on Newton-Raphson algorithm	Experimental verification, full filed deformation measurement	Without any loss of measurement accuracy, the calculation speed is increased by 120~200 times
	Pan et al. [151] (2013)	DIC template matching based on IC-GN	Numerical tests and experimental verification	The proposed IC-GN is 3~5 times faster than Newton-Raphson
	Zhang et al. [145] (2016)	Integrate two efficient subpixel level motion extraction algorithms	Experimental verification, Taylor approximation refinement algorithm and the localization refinement algorithm, dynamic vibration analysis	In the case of similar accuracy, it is at least 5 times faster than the traditional UCC method (RMS error 0.75%)
	Xu et al. [157] (2019)	Fuse the vision-based displacement measurement with acceleration data	Field test, short-span railway bridge, displacement ~ time history curve	The RMS of measurement noise at the camera-to-target distance of 6.9 m is less than 0.2 mm

4.5. Other Impacts

An SHM system requires long-term stability in order to assess the state and performance of a structure. The displacement measurement methods in the literature mainly focus on short-term measurements of up to several hours. For long-term measurements of up to several months or even years, the uncertainty and reliability of the computer vision system still need further research. In addition, data acquisition, data transmission and data processing in computer vision systems are new challenges and require highly specialized personnel during and after equipment installation and maintenance, which will require new resources.

When multiple vision-based displacement measurement subsystems need to work together or to measure the displacement of multiple measurement points at the same time [158], the problem of time synchronization of multiple cameras must be solved. Luo et al. [159] developed a vision-based synchronization system using master/slave systems for wireless data communication in order to simultaneously measure multiple points of the structure. Fukuda et al. [2] developed a time synchronization system that connected multiple displacement measurement subsystems using a local area network, enabling computers to communicate with each other using the TCP/IP protocol. Dong et al. [74] synchronized all cameras and potentiometers using an NI Multifunction I/O Device to ensure that multiple cameras and sensors worked synchronously.

5. Conclusions and Prospects

This paper briefly describes the composition of computer vision monitoring systems, introduces the basic monitoring process and methods, and pays special attention to the problems and solutions encountered in the application of computer vision in the field environment. From the examined articles, the following main conclusions can be made:

(1) At present, the main application of computer vision in the field of SHM is still focused on the measurement of displacement time history curves of scale models under static and dynamic loading in controlled conditions and for short terms.

(2) A large number of experimental tests and short-term field tests promote the formation of the basic framework of computer vision deformation monitoring systems, and existing research has focused on improving the applicability and stability of image processing algorithms.

(3) Structural deformation monitoring systems based on computer vision have had some solutions to cope with individual external influences (such as target installation difficulty, illuminate change, camera movement and climate transformation). The accuracy and reliability of computer vision-based structural deformation monitoring has made great progress and is gradually approaching practical long-term monitoring.

It has been more than 30 years since computer vision was first applied to civil engineering structural measurement. Vision based sensors have made great progress and achievements in technology, but they still face some limitations and challenges. In the future, we need to do more in these aspects: (1) devise simpler programs or devices to promote their long-term applications in practical engineering; and (2) uncertainty evaluation of vision sensors in long-term applications.

References

1. Feng, D.M.; Feng, M.Q.; Ozer, E.; Fukuda, Y. A vision-based sensor for noncontact structural displacement measurement. *Sensors* **2015**, *15*, 16557–16575. [CrossRef] [PubMed]
2. Fukuda, Y.; Feng, M.Q.; Shinozuka, M. Cost-effective vision-based system for monitoring dynamic response of civil engineering structures. *Struct. Control Health Monit.* **2010**, *17*, 918–936. [CrossRef]
3. Feng, D.M.; Feng, M.Q. Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection—A review. *Eng. Struct.* **2018**, *156*, 105–117. [CrossRef]
4. Bastianini, F.; Corradi, M.; Borri, A.; Tommaso, A.D. Retrofit and monitoring of an historical building using "Smart" CFRP with embedded fibre optic Brillouin sensors. *Constr. Build. Mater.* **2005**, *19*, 525–535. [CrossRef]
5. Chan, T.H.T.; Yu, L.; Tam, H.Y.; Ni, Y.-Q.; Liu, S.Y.; Chung, W.H.; Cheng, L.K. Fiber Bragg grating sensors for structural health monitoring of Tsing Ma bridge: Background and experimental observation. *Eng. Struct.* **2006**, *28*, 648–659. [CrossRef]
6. He, J.; Zhou, Z.; Jinping, O. Optic fiber sensor-based smart bridge cable with functionality of self-sensing. *Mech. Syst. Signal Process.* **2013**, *35*, 84–94. [CrossRef]
7. Li, H.; Ou, J.; Zhou, Z. Applications of optical fibre Bragg gratings sensing technology-based smart stay cables. *Opt. Lasers Eng.* **2009**, *47*, 1077–1084. [CrossRef]
8. Metje, N.; Chapman, D.; Rogers, C.; Henderson, P.; Beth, M. An Optical Fiber Sensor System for Remote Displacement Monitoring of Structures—Prototype Tests in the Laboratory. *Struct. Health Monit.* **2008**, *7*, 51–63. [CrossRef]
9. Rodrigues, C.; Cavadas, F.; Félix, C.; Figueiras, J. FBG based strain monitoring in the rehabilitation of a centenary metallic bridge. *Eng. Struct.* **2012**, *44*, 281–290. [CrossRef]
10. Hester, D.; Brownjohn, J.; Bocian, M.; Xu, Y. Low cost bridge load test: Calculating bridge displacement from acceleration for load assessment calculations. *Eng. Struct.* **2017**, *143*, 358–374. [CrossRef]
11. Gindy, M.; Nassif, H.H.; Velde, J. Bridge Displacement Estimates from Measured Acceleration Records. *Transp. Res. Rec. J. Transp. Res. Board* **2007**, *2028*, 136–145. [CrossRef]
12. Çelebi, M. GPS in dynamic monitoring of long-period structures. *Soil Dyn. Earthq. Eng.* **2000**, *20*, 477–483. [CrossRef]
13. Nakamura, S.-I. GPS Measurement of Wind-Induced Suspension Bridge Girder Displacements. *J. Struct. Eng.* **2000**, *126*, 1413–1419. [CrossRef]
14. Xu, L.; Guo, J.J.; Jiang, J.J. Time–frequency analysis of a suspension bridge based on GPS. *J. Sound Vib.* **2002**, *254*, 105–116. [CrossRef]
15. Garg, P.; Moreu, F.; Ozdagli, A.; Taha, M.R.; Mascareñas, D. Noncontact Dynamic Displacement Measurement of Structures Using a Moving Laser Doppler Vibrometer. *J. Bridg. Eng.* **2019**, *24*, 04019089. [CrossRef]
16. Teskey, R.S.; Radovanovic, W.F. Dynamic monitoring of deforming structures: Gps versus robotic tacheometry systems. In Proceedings of the International Symposium on Deformation Measurements, Orange, CA, USA, 19–22 March 2001.
17. Pieraccini, M.; Parrini, F.; Fratini, M.; Atzeni, C.; Spinelli, P.; Micheloni, M. Static and dynamic testing of bridges through microwave interferometry. *NDT E Int.* **2007**, *40*, 208–214. [CrossRef]
18. Chang, C.-C.; Xiao, X.H. Three-Dimensional Structural Translation and Rotation Measurement Using Monocular Videogrammetry. *J. Eng. Mech.* **2010**, *136*, 840–848. [CrossRef]
19. Fukuda, Y.; Feng, M.Q.; Narita, Y.; Kaneko, S.; Tanaka, T. Vision-based displacement sensor for monitoring dynamic response using robust object search algorithm. *IEEE Sens. J.* **2013**, *13*, 4725–4732. [CrossRef]
20. Kohut, P.; Holak, K.; Uhl, T.; Ortyl, Ł.; Owerko, T.; Kuras, P.; Kocierz, R. Monitoring of a civil structure's state based on noncontact measurements. *Struct. Health Monit.* **2013**, *12*, 411–429. [CrossRef]
21. Ribeiro, D.; Calçada, R.; Ferreira, J.; Martins, T. Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system. *Eng. Struct.* **2014**, *75*, 164–180. [CrossRef]
22. Casciati, F.; Fuggini, C. Monitoring a steel building using GPS sensors. *Smart Struct. Syst.* **2011**, *7*, 349–363. [CrossRef]
23. Xu, Y.; Brownjohn, J.M.W.; Hester, D.; Koo, K.Y. Long-span bridges: Enhanced data fusion of GPS displacement and deck accelerations. *Eng. Struct.* **2017**, *147*, 639–651. [CrossRef]
24. Mehrabi, A.B. In-Service Evaluation of Cable-Stayed Bridges, Overview of Available Methods and Findings. *J. Bridg. Eng.* **2006**, *11*, 716–724. [CrossRef]
25. Nassif, H.H.; Gindy, M.; Davis, J. Comparison of laser Doppler vibrometer with contact sensors for monitoring bridge deflection and vibration. *NDT E Int.* **2005**, *38*, 213–218. [CrossRef]
26. Casciati, F.; Wu, L.J. Local positioning accuracy of laser sensors for structural health monitoring. *Struct. Control Health Monit.* **2012**, *20*, 728–739. [CrossRef]
27. Dong, C.Z.; Bas, S.; Catbas, N. A completely non-contact recognition system for bridge unit influence line using portable cameras and computer vision. *Smart Struct Syst.* **2019**, *24*, 617–630.
28. Zaurin, R.; Catbas, F.N. Structural health monitoring using video stream, influence lines, and statistical analysis. *Struct. Health Monit.* **2010**, *10*, 309–332. [CrossRef]
29. Khuc, T.; Catbas, F.N. Structural Identification Using Computer Vision–Based Bridge Health Monitoring. *J. Struct. Eng.* **2018**, *144*, 04017202. [CrossRef]

30. Zhang, L.H.; Jin, Y.J.; Lin, L.; Li, J.J.; Du, Y.A. The comparison of ccd and cmos image sensors. In Proceedings of the International Conference on Optical Instruments and Technology—Advanced Sensor Technologies and Applications, Beijing, China, 16–19 November 2009.

31. Bigas, M.; Cabruja, E.; Forest, J.; Salvi, J. Review of CMOS image sensors. *Microelectron. J.* **2006**, *37*, 433–451. [CrossRef]

32. Khuc, T.; Catbas, F.N. Computer vision-based displacement and vibration monitoring without using physical target on structures. *Struct. Infrastruct. Eng.* **2017**, *13*, 505–516. [CrossRef]

33. Lee, J.; Lee, K.-C.; Cho, S.; Sim, S.-H. Computer Vision-Based Structural Displacement Measurement Robust to Light-Induced Image Degradation for In-Service Bridges. *Sensors* **2017**, *17*, 2317. [CrossRef] [PubMed]

34. Lee, J.-H.; Ho, H.-N.; Shinozuka, M.; Lee, J.-J. An advanced vision-based system for real-time displacement measurement of high-rise buildings. *Smart Mater. Struct.* **2012**, *21*, 125019. [CrossRef]

35. Ye, X.; Dong, C. Review of computer vision-based structural displacement monitoring. *China J. Highw. Transp.* **2019**, *32*, 21–39.

36. Ji, Y.F. A computer vision-based approach for structural displacement measurement. In Proceedings of the Conference on Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2010, San Diego, CA, USA, 7–11 March 2010; Spie-Int Soc Optical Engineering: Bellingham, DC, USA, 2010.

37. Shariati, A.; Schumacher, T. Eulerian-based virtual visual sensors to measure dynamic displacements of structures. *Struct. Control Health Monit.* **2017**, *24*, e1977. [CrossRef]

38. Yoon, H.; Elanwar, H.; Choi, H.; Golparvar-Fard, M.; Spencer, B.F., Jr. Target-free approach for vision-based structural system identification using consumer-grade cameras. *Struct. Control Health Monit.* **2016**, *23*, 1405–1416. [CrossRef]

39. Kuddus, M.A.; Li, J.; Hao, H.; Li, C.; Bi, K. Target-free vision-based technique for vibration measurements of structures subjected to out-of-plane movements. *Eng. Struct.* **2019**, *190*, 210–222. [CrossRef]

40. Lepetit, V.; Fua, P. Monocular Model-Based 3D Tracking of Rigid Objects: A Survey. *Found. Trends Comput. Graph. Vis.* **2005**, *1*, 89. [CrossRef]

41. Chang, C.-C.; Ji, Y.F. Flexible Videogrammetric Technique for Three-Dimensional Structural Vibration Measurement. *J. Eng. Mech.* **2007**, *133*, 656–664. [CrossRef]

42. Dong, C.-Z.; Catbas, F.N. A non-target structural displacement measurement method using advanced feature matching strategy. *Adv. Struct. Eng.* **2019**, *22*, 3461–3472. [CrossRef]

43. Zhang, Z. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330–1334. [CrossRef]

44. Heikkila, J. Geometric camera calibration using circular control points. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1066–1077. [CrossRef]

45. Fleet, D.J.; Jepson, A.D. Computation of component image velocity from local phase information. *Int. J. Comput. Vis.* **1990**, *5*, 77–104. [CrossRef]

46. Sturm, P.F.; Maybank, S.J. On plane-based camera calibration: A general algorithm, singularities, applications. In Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'99), Fort Collins, CO, USA, 23–25 June 1999; IEEE: Piscataway, NJ, USA, 1999.

47. Xu, Y.; Zhang, J.; Brownjohn, J. An accurate and distraction-free vision-based structural displacement measurement method integrating Siamese network based tracker and correlation-based template matching. *Measurement* **2021**, *179*, 109506. [CrossRef]

48. Xu, Y.; Brownjohn, J.M.W.; Hester, D.; Koo, K.-Y. Dynamic displacement measurement of a long-span bridge using vision-based system. In Proceedings of the 8th European Workshop on Structural Health Monitoring, EWSHM 2016, Bilbao, Spain, 5–8 July 2016.

49. Ye, X.W.; Yi, T.-H.; Dong, C.Z.; Liu, T. Vision-based structural displacement measurement: System performance evaluation and influence factor analysis. *Measurement* **2016**, *88*, 372–384. [CrossRef]

50. Ye, X.W.; Jin, T.; Ang, P.-P.; Bian, X.-C.; Chen, Y.-M. Computer vision-based monitoring of the 3-d structural deformation of an ancient structure induced by shield tunneling construction. *Struct. Control Health Monit.* **2021**, *28*, e2702. [CrossRef]

51. Feng, D.; Feng, M.Q. Experimental validation of cost-effective vision-based structural health monitoring. *Mech. Syst. Signal Process.* **2017**, *88*, 199–211. [CrossRef]

52. Song, Y.-Z.; Bowen, C.R.; Kim, A.H.; Nassehi, A.; Padget, J.; Gathercole, N. Virtual visual sensors and their application in structural health monitoring. *Struct. Health Monit.* **2014**, *13*, 251–264. [CrossRef]

53. Ye, X.W.; Ni, Y.-Q.; Wai, T.; Wong, K.; Zhang, X.; Xu, F. A vision-based system for dynamic displacement measurement of long-span bridges: Algorithm and verification. *Smart Struct. Syst.* **2013**, *12*, 363–379. [CrossRef]

54. Wahbeh, A.M.; Caffrey, J.P.; Masri, S.F. A vision-based approach for the direct measurement of displacements in vibrating systems. *Smart Mater. Struct.* **2003**, *12*, 785–794. [CrossRef]

55. Ji, Y.F.; Chang, C.C. Nontarget Image-Based Technique for Small Cable Vibration Measurement. *J. Bridg. Eng.* **2008**, *13*, 34–42. [CrossRef]

56. Ghosal, S.; Mehrotra, R. Orthogonal moment operators for subpixel edge detection. *Pattern Recognit.* **1993**, *26*, 295–306. [CrossRef]

57. Sobel, I. Neighborhood coding of binary images for fast contour following and general binary array processing. *Comput. Graph. Image Process.* **1978**, *8*, 127–135. [CrossRef]

58. Marr, D.; Hildreth, E. Theory of edge detection. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **1980**, *207*, 187–217.

59. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *8*, 679–698. [CrossRef]

60. Ballard, D. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognit.* **1981**, *13*, 111–122. [CrossRef]

61. Ji, Y.F.; Chang, C.C. Nontarget Stereo Vision Technique for Spatiotemporal Response Measurement of Line-Like Structures. *J. Eng. Mech.* **2008**, *134*, 466–474. [CrossRef]
62. Tian, Y.D.; Zhang, J.; Yu, S.S. Vision-based structural scaling factor and flexibility identification through mobile impact testing. *Mech. Syst. Signal Process.* **2019**, *122*, 387–402. [CrossRef]
63. Xu, Y.; Brownjohn, J.; Kong, D. A non-contact vision-based system for multipoint displacement monitoring in a cable-stayed footbridge. *Struct. Control Health Monit.* **2018**, *25*, e2155. [CrossRef]
64. Khuc, T.; Catbas, F.N. Completely contactless structural health monitoring of real-life structures using cameras and computer vision. *Struct. Control Health Monit.* **2017**, *24*, e1852. [CrossRef]
65. Harris, C.G.; Stephens, M. A Combined Corner and Edge Detector. In Proceedings of the Alvey Vision Conference, Manchester, UK, 31 August–2 September 1988. [CrossRef]
66. Jianbo, S.; Tomasi, C. Good features to track. In Proceedings of the 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat No94CH3405-8), Seattle, WA, USA, 21–23 June 1994; pp. 593–600.
67. Gao, K.; Lin, S.X.; Zhang, Y.D.; Tang, S.; Ren, H.M. Attention model based sift keypoints filtration for image retrieval. In Proceedings of the 7th IEEE/ACIS International Conference on Computer and Information Science in Conjunction with 2nd IEEE/ACIS International Workshop on e-Activity, Portland, OR, USA, 14–16 May 2008.
68. Bay, H.; Tuytelaars, T.; Van Gool, L. Surf: Speeded up robust features. In *Computer Vision—ECCV 2006, Proceedings of the 9th European Conference on Computer Vision, Graz, Austria, 7–13 May 2006*; Leonardis, A., Bischof, H., Pinz, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 404–417.
69. Calonder, M.; Lepetit, V.; Strecha, C.; Fua, P. Brief: Binary robust independent elementary features. In *Computer Vision—ECCV 2010, Proceedings of the 11th European Conference on Computer Vision, Crete, Greece, 5–11 September 2010*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 778–792.
70. Leutenegger, S.; Chli, M.; Siegwart, R.Y. Brisk: Binary robust invariant scalable keypoints. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Barcelona, Spain, 6–13 November 2011.
71. Alahi, A.; Ortiz, R.; Vandergheynst, P. Freak: Fast retina keypoint. 2012 ieee conference on computer vision and pattern recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; IEEE: New York, NY, USA, 2012; pp. 510–517.
72. Dong, C.-Z.; Catbas, F.N. A review of computer vision–based structural health monitoring at local and global levels. *Struct. Health Monit.* **2020**, *20*, 692–743. [CrossRef]
73. Sun, D.Q.; Roth, S.; Black, M.J. Secrets of optical flow estimation and their principles. In Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition, Los Alamitos, CA, USA, 13–18 June 2010; pp. 2432–2439.
74. Dong, C.-Z.; Celik, O.; Catbas, F.N. Marker-free monitoring of the grandstand structures and modal identification using computer vision methods. *Struct. Health Monit.* **2019**, *18*, 1491–1509. [CrossRef]
75. Chen, Z.; Gao, M.; Shen, Y. The current situation and trends of optical flow estimation. *J. Image Graph.* **2002**, *7*, 434–439.
76. Lucas, B.D.; Kanade, T. An iterative image registration technique with an application to stereo vision (darpa). In Proceedings of the 7th International Joint Conference on Artificial Intelligence, Vancouver, BC, Canada, 24–28 August 1981; pp. 674–679.
77. Celik, O.; Dong, C.-Z.; Catbas, F.N. A computer vision approach for the load time history estimation of lively individuals and crowds. *Comput. Struct.* **2018**, *200*, 32–52. [CrossRef]
78. Horn, B.K.; Schunck, B.G. Determining optical flow. *Artif. Intell.* **1981**, *17*, 185–203. [CrossRef]
79. Farneback, G. Two-frame motion estimation based on polynomial expansion. In *Image Analysis, Proceedings of the 13th Scandinavian Conference, SCIA 2003 Halmstad, Sweden, 29 June–2 July 2003*; Lecture Notes in Computer Science; Bigun, J., Gustavsson, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2003; pp. 363–370.
80. Liu, B.; Zaccarin, A. New fast algorithms for the estimation of block motion vectors. *IEEE Trans. Circuits Syst. Video Technol.* **1993**, *3*, 148–157. [CrossRef]
81. Wadhwa, N.; Rubinstein, M.; Durand, F.; Freeman, W.T. Phase-based video motion processing. *ACM Trans. Graph.* **2013**, *32*, 80. [CrossRef]
82. Chen, J.G.; Wadhwa, N.; Cha, Y.-J.; Durand, F.; Freeman, W.T.; Buyukozturk, O. Modal identification of simple structures with high-speed video using motion magnification. *J. Sound Vib.* **2015**, *345*, 58–71. [CrossRef]
83. Kim, S.W.; Jeon, B.-G.; Cheung, J.-H.; Kim, S.-D.; Park, J.-B. Stay cable tension estimation using a vision-based monitoring system under various weather conditions. *J. Civ. Struct. Health Monit.* **2017**, *7*, 343–357. [CrossRef]
84. Pan, B.; Qian, K.M.; Xie, H.M.; Asundi, A. Two-dimensional digital image correlation for in-plane displacement and strain measurement: A review. *Meas. Sci. Technol.* **2009**, *20*, 17. [CrossRef]
85. Fayyad, T.M.; Lees, J.M. Experimental investigation of crack propagation and crack branching in lightly reinforced concrete beams using digital image correlation. *Eng. Fract. Mech.* **2017**, *182*, 487–505. [CrossRef]
86. Ghiassi, B.; Xavier, J.; Oliveira, D.V.; Lourenço, P.B. Application of digital image correlation in investigating the bond between FRP and masonry. *Compos. Struct.* **2013**, *106*, 340–349. [CrossRef]
87. Guizar-Sicairos, M.; Thurman, S.T.; Fienup, J.R. Efficient subpixel image registration algorithms. *Opt. Lett.* **2008**, *33*, 156–158. [CrossRef] [PubMed]
88. Swain, M.J.; Ballard, D.H. Color indexing. *Int. J. Comput. Vis.* **1991**, *7*, 11–32. [CrossRef]

89. Jing, H.; Kumar, S.R.; Mitra, M.; Wei-Jing, Z.; Zabih, R. Image indexing using color correlograms. In Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat No97CB36082), San Juan, PR, USA, 17–19 June 1997; pp. 762–768.

90. Cigada, A.; Mazzoleni, P.; Zappa, E. Vibration Monitoring of Multiple Bridge Points by Means of a Unique Vision-Based Measuring System. *Exp. Mech.* **2014**, *54*, 255–271. [CrossRef]

91. Stephen, G.A.; Brownjohn, J.M.W.; Taylor, C.A. Measurements of static and dynamic displacement from visual monitoring of the Humber Bridge. *Eng. Struct.* **1993**, *15*, 197–208. [CrossRef]

92. Macdonald, J.H.G.; Dagless, E.L.; Thomas, B.T.; Taylor, C.A. Dynamic measurements of the second severn crossing. *Proc. Inst. Civ. Eng. Transp.* **1997**, *123*, 241–248. [CrossRef]

93. Tian, L.; Pan, B. Remote Bridge Deflection Measurement Using an Advanced Video Deflectometer and Actively Illuminated LED Targets. *Sensors* **2016**, *16*, 1344. [CrossRef]

94. Kim, H.; Shin, S. Reliability verification of a vision-based dynamic displacement measurement for system identification. *J. Wind Eng. Ind. Aerodyn.* **2019**, *191*, 22–31. [CrossRef]

95. Dong, C.Z.; Ye, X.W.; Jin, T. Identification of structural dynamic characteristics based on machine vision technology. *Measurement* **2018**, *126*, 405–416. [CrossRef]

96. Lee, J.J.; Shinozuka, M. A vision-based system for remote sensing of bridge displacement. *NDT E Int.* **2006**, *39*, 425–431. [CrossRef]

97. Ramana, L.; Choi, W.; Cha, Y.-J. Fully automated vision-based loosened bolt detection using the Viola–Jones algorithm. *Struct. Health Monit.* **2018**, *18*, 422–434. [CrossRef]

98. Sun, J.H.; Xie, Y.X.; Cheng, X.Q. A Fast Bolt-Loosening Detection Method of Running Train's Key Components Based on Binocular Vision. *IEEE Access* **2019**, *7*, 32227–32239. [CrossRef]

99. Wang, C.Y.; Wang, N.; Ho, S.-C.; Chen, X.M.; Song, G.B. Design of a New Vision-Based Method for the Bolts Looseness Detection in Flange Connections. *IEEE Trans. Ind. Electron.* **2020**, *67*, 1366–1375. [CrossRef]

100. Cha, Y.-J.; Choi, W.; Suh, G.; Mahmoudkhani, S.; Büyüköztürk, O. Autonomous Structural Visual Inspection Using Region-Based Deep Learning for Detecting Multiple Damage Types. *Comput. Civ. Infrastruct. Eng.* **2017**, *33*, 731–747. [CrossRef]

101. Pan, X.; Yang, T.Y. Postdisaster image-based damage detection and repair cost estimation of reinforced concrete buildings using dual convolutional neural networks. *Comput. Civ. Infrastruct. Eng.* **2020**, *35*, 495–510. [CrossRef]

102. Liang, X. Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization. *Comput. Civ. Infrastruct. Eng.* **2019**, *34*, 415–430. [CrossRef]

103. Muhammad, K.; Ahmad, J.; Lv, Z.; Bellavista, P.; Yang, P.; Baik, S.W. Efficient Deep CNN-Based Fire Detection and Localization in Video Surveillance Applications. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *49*, 1419–1434. [CrossRef]

104. He, Y.; Zhang, L.; Chen, Z.; Li, C.Y. A framework of structural damage detection for civil structures using a combined multi-scale convolutional neural network and echo state network. *Eng. Comput.* **2022**, 1–19. [CrossRef]

105. Kim, S.W.; Kim, N.-S. Dynamic characteristics of suspension bridge hanger cables using digital image processing. *NDT E Int.* **2013**, *59*, 25–33. [CrossRef]

106. Kim, S.W.; Jeon, B.-G.; Kim, N.-S.; Park, J.-C. Vision-based monitoring system for evaluating cable tensile forces on a cable-stayed bridge. *Struct. Health Monit.* **2013**, *12*, 440–456. [CrossRef]

107. Ye, X.W.; Dong, C.Z.; Liu, T. Force monitoring of steel cables using vision-based sensing technology: Methodology and experimental verification. *Smart Struct. Syst.* **2016**, *18*, 585–599. [CrossRef]

108. Chen, C.-C.; Wu, W.-H.; Tseng, H.-Z.; Chen, C.-H.; Lai, G.L. Application of digital photogrammetry techniques in identifying the mode shape ratios of stay cables with multiple camcorders. *Measurement* **2015**, *75*, 134–146. [CrossRef]

109. Park, J.-W.; Lee, J.-J.; Jung, H.-J.; Myung, H. Vision-based displacement measurement method for high-rise building structures using partitioning approach. *NDT E Int.* **2010**, *43*, 642–647. [CrossRef]

110. Feng, D.; Scarangello, T.; Feng, M.Q.; Ye, Q. Cable tension force estimate using novel noncontact vision-based sensor. *Measurement* **2017**, *99*, 44–52. [CrossRef]

111. Feng, M.Q.; Fukuda, Y.; Feng, D.; Mizuta, M. Nontarget Vision Sensor for Remote Measurement of Bridge Dynamic Response. *J. Bridge Eng.* **2015**, *20*, 04015023. [CrossRef]

112. Brownjohn, J.M.W.; Xu, Y.; Hester, D. Vision-Based Bridge Deformation Monitoring. *Front. Built Environ.* **2017**, *3*, 23. [CrossRef]

113. Li, J.-C.; Yuan, B.-Z. Using vision technique for the bridge deformation detection. In Proceedings of the ICASSP '88: 1988 International Conference on Acoustics, Speech, and Signal Processing, New York, NY, USA, 11–14 April 1988; IEEE: New York, NY, USA, 1988.

114. Feng, D.; Feng, M.Q. Vision-based multipoint displacement measurement for structural health monitoring. *Struct. Control Health Monit.* **2016**, *23*, 876–890. [CrossRef]

115. Feng, D.; Feng, M.Q. Model Updating of Railway Bridge Using In Situ Dynamic Displacement Measurement under Trainloads. *J. Bridg. Eng.* **2015**, *20*, 04015019. [CrossRef]

116. Lee, J.-J.; Ho, H.-N.; Lee, J.H. A Vision-Based Dynamic Rotational Angle Measurement System for Large Civil Structures. *Sensors* **2012**, *12*, 7326–7336. [CrossRef]

117. Chen, Z.C.; Li, H.; Bao, Y.Q.; Li, N.; Jin, Y. Identification of spatio-temporal distribution of vehicle loads on long-span bridges using computer vision technology. *Struct. Control Health Monit.* **2016**, *23*, 517–534. [CrossRef]

118. Yoon, H.; Shin, J.; Spencer, B.F. Structural Displacement Measurement Using an Unmanned Aerial System. *Comput. Civ. Infrastruct. Eng.* **2018**, *33*, 183–192. [CrossRef]

119. Ma, S.P.; Pang, J.Z.; Ma, Q.W. The systematic error in digital image correlation induced by self-heating of a digital camera. *Meas. Sci. Technol.* **2012**, *23*, 7. [CrossRef]

120. Ehrhart, M.; Lienhart, W. Image-based dynamic deformation monitoring of civil engineering structures from long ranges. In *Image Processing: Machine Vision Applications viii. Proceedings of Spie. 9405*; Lam, E.Y., Niel, K.S., Eds.; Spie-Int Soc Optical Engineering: Bellingham, DC, USA, 2015.

121. Ye, X.W.; Yi, T.-H.; Dong, C.Z.; Liu, T.; Bai, H. Multi-point displacement monitoring of bridges using a vision-based approach. *Wind Struct. Int. J.* **2015**, *20*, 315–326. [CrossRef]

122. Anantrasirichai, N.; Achim, A.; Kingsbury, N.G.; Bull, D.R. Atmospheric Turbulence Mitigation Using Complex Wavelet-Based Fusion. *IEEE Trans. Image Process.* **2013**, *22*, 2398–2408. [CrossRef]

123. Shimizu, M.; Yoshimura, S.; Tanaka, M.; Okutomi, M. Super-resolution from image sequence under influence of hot-air optical turbulence. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008.

124. Zhu, X.; Milanfar, P. Removing Atmospheric Turbulence via Space-Invariant Deconvolution. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 157–170. [CrossRef]

125. Luo, L.X.; Feng, M.Q.; Wu, J. A comprehensive alleviation technique for optical-turbulence-induced errors in vision-based displacement measurement. *Struct. Control Health Monit.* **2020**, *27*, 16. [CrossRef]

126. Luo, L.X.; Feng, M.Q.; Wu, J.P.; Bi, L.Z. Modeling and Detection of Heat Haze in Computer Vision Based Displacement Measurement. *Measurement* **2021**, *182*, 109772. [CrossRef]

127. Owens, J.C. Optical Refractive Index of Air: Dependence on Pressure, Temperature and Composition. *Appl. Opt.* **1967**, *6*, 51–59. [CrossRef]

128. Yoneyama, S.; Ueda, H. Bridge Deflection Measurement Using Digital Image Correlation with Camera Movement Correction. *Mater. Trans.* **2012**, *53*, 285–290. [CrossRef]

129. Hoskere, V.; Park, J.-W.; Yoon, H.; Spencer, B.F., Jr. Vision-Based Modal Survey of Civil Infrastructure Using Unmanned Aerial Vehicles. *J. Struct. Eng.* **2019**, *145*, 04019062. [CrossRef]

130. Casciati, F.; Casciati, S.; Colnaghi, A.; Faravelli, L.; Rosadini, L.; Zhu, S. Vision-based support in the characterization of superelastic u-shaped sma elements. *Smart Struct. Syst.* **2019**, *24*, 641–648.

131. Ribeiro, D.; Santos, R.; Cabral, R.; Saramago, G.; Montenegro, P.; Carvalho, H.; Correia, J.; Calçada, R. Non-contact structural displacement measurement using Unmanned Aerial Vehicles and video-based systems. *Mech. Syst. Signal Process.* **2021**, *160*, 107869. [CrossRef]

132. Lee, J.; Lee, K.-C.; Jeong, S.; Lee, Y.-J.; Sim, S.-H. Long-term displacement measurement of full-scale bridges using camera ego-motion compensation. *Mech. Syst. Signal Process.* **2020**, *140*, 106651. [CrossRef]

133. Chen, J.G.; Davis, A.; Wadhwa, N.; Durand, F.; Freeman, W.T.; Büyüköztürk, O. Video Camera–Based Vibration Measurement for Civil Infrastructure Applications. *J. Infrastruct. Syst.* **2017**, *23*, 11. [CrossRef]

134. Khaloo, A.; Lattanzi, D. Pixel-wise structural motion tracking from rectified repurposed videos. *Struct. Control Health Monit.* **2017**, *24*, e2009. [CrossRef]

135. Rousseeuw, P.J.; Hubert, M. Robust statistics for outlier detection. *Wiley Interdiscip. Rev-Data Min. Knowl. Discov.* **2011**, *1*, 73–79. [CrossRef]

136. Spencer, B.F., Jr.; Hoskere, V.; Narazaki, Y. Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring. *Engineering* **2019**, *5*, 199–222. [CrossRef]

137. Li, Y.; Takauji, H.; Ohmura, I.; Kaneko, S.; Tanaka, T. Robust Focusing using Orientation Code Matching. *J. Jpn. Soc. Precis. Eng.* **2009**, *75*, 650–656. [CrossRef]

138. Ullah, F.; Kaneko, S.; Igarashi, S. Orientation code matching for robust object search. *IEICE Trans. Inf. Syst.* **2001**, *E84D*, 999–1006.

139. Luo, L.; Feng, M.Q. Edge-Enhanced Matching for Gradient-Based Computer Vision Displacement Measurement. *Comput. Civ. Infrastruct. Eng.* **2018**, *33*, 1019–1040. [CrossRef]

140. Luo, L.; Feng, M.Q.; Wu, Z.Y. Robust vision sensor for multi-point displacement monitoring of bridges in the field. *Eng. Struct.* **2018**, *163*, 255–266. [CrossRef]

141. Lichao, Z.; Gonzalez-Garcia, A.; Van De Weijer, J.; Danelljan, M.; Khan, F.S. Learning the model update for siamese trackers. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27–28 October 2019; pp. 4009–4018.

142. Dong, C.-Z.; Celik, O.; Catbas, F.N.; Obrien, E.; Taylor, S. A Robust Vision-Based Method for Displacement Measurement under Adverse Environmental Factors Using Spatio-Temporal Context Learning and Taylor Approximation. *Sensors* **2019**, *19*, 3197. [CrossRef] [PubMed]

143. Javh, J.; Slavič, J.; Boltežar, M. High frequency modal identification on noisy high-speed camera data. *Mech. Syst. Signal Process.* **2018**, *98*, 344–351. [CrossRef]

144. Lecompte, D.; Smits, A.; Bossuyt, S.; Sol, H.; Vantomme, J.; Van Hemelrijck, D.; Habraken, A.M. Quality assessment of speckle patterns for digital image correlation. *Opt. Lasers Eng.* **2006**, *44*, 1132–1145. [CrossRef]

145. Zhang, D.; Guo, J.; Lei, X.; Zhu, C. A High-Speed Vision-Based Sensor for Dynamic Vibration Analysis Using Fast Motion Extraction Algorithms. *Sensors* **2016**, *16*, 572. [CrossRef]
146. Guo, J.; Zhu, C. Dynamic displacement measurement of large-scale structures based on the Lucas–Kanade template tracking algorithm. *Mech. Syst. Signal Process.* **2016**, *66–67*, 425–436. [CrossRef]
147. Pan, B.; Xie, H.M.; Xu, B.Q.; Dai, F.L. Performance of sub-pixel registration algorithms in digital image correlation. *Meas Sci Technol.* **2006**, *17*, 1615–1621.
148. Macvicar-Whelan, P.J.; Binford, T.O. Line finding with subpixel precision. *Proc. SPIE—Int. Soc. Opt. Eng.* **1981**, *281*, 211–216.
149. Jensen, K.; Anastassiou, D. Subpixel edge localization and the interpolation of still images. *IEEE Trans. Image Process.* **1995**, *4*, 285–295. [CrossRef]
150. Bruck, H.A.; Mcneill, S.R.; Sutton, M.A.; Peters, W.H. Digital image correlation using newton-raphson method of partial differential correction displacement measurement. *Exp. Mech.* **1989**, *29*, 261–267. [CrossRef]
151. Pan, B.; Li, K.; Tong, W. Fast, Robust and Accurate Digital Image Correlation Calculation Without Redundant Computations. *Exp. Mech.* **2013**, *53*, 1277–1289. [CrossRef]
152. Qu, Y.D.; Cui, C.S.; Chen, S.B.; Li, J.Q. A fast subpixel edge detection method using sobel-zernike moments operator. *Image Vis Comput.* **2005**, *23*, 11–17.
153. Mas, D.; Perez, J.; Ferrer, B.; Espinosa, J. Realistic limits for subpixel movement detection. *Appl. Opt.* **2016**, *55*, 4974–4979. [CrossRef] [PubMed]
154. Khuc, T.; Nguyen, T.A.; Dao, H.; Catbas, F.N. Swaying displacement measurement for structural monitoring using computer vision and an unmanned aerial vehicle. *Measurement* **2020**, *159*, 107769. [CrossRef]
155. Shan, B.H.; Zheng, S.J.; Ou, J.P. Free vibration monitoring experiment of a stayed-cable model based on stereovision. *Measurement* **2015**, *76*, 228–239. [CrossRef]
156. Pan, B.; Li, K. A fast digital image correlation method for deformation measurement. *Opt. Lasers Eng.* **2011**, *49*, 841–847. [CrossRef]
157. Xu, Y.; Brownjohn, J.M.W.; Huseynov, F. Accurate Deformation Monitoring on Bridge Structures Using a Cost-Effective Sensing System Combined with a Camera and Accelerometers: Case Study. *J. Bridg. Eng.* **2019**, *24*, 05018014. [CrossRef]
158. Lee, J.J.; Shinozuka, M. Real-Time Displacement Measurement of a Flexible Bridge Using Digital Image Processing Techniques. *Exp. Mech.* **2006**, *46*, 105–114. [CrossRef]
159. Lee, J.J.; Fukuda, Y.; Shinozuka, M.; Cho, S.; Yun, C.-B. Development and application of a vision-based displacement measurement system for structural health monitoring of civil structures. *Smart Struct. Syst.* **2007**, *3*, 373–384. [CrossRef]