



Monocular Vision-Based Depth Displacement Measurement of Base Isolators

Weiquan Luo^a, Yu Dang*

School of Civil Engineering, Lanzhou University of Technology, Lanzhou, Gansu, 730050, China

^ae-mail: wq.luo@foxmail.com

*Corresponding author's e-mail address: 601363791@qq.com

Abstract. In response to the simultaneous movement of seismic isolators and cameras during earthquakes, this paper proposes a three-dimensional displacement measurement method based on monocular vision. This method involves placing a chessboard target on the upper and lower supports of the isolator and using a camera to monitor the target. The system identifies image feature points and calculates the relative position between the camera and the target using corner detection and the PnP algorithm, thereby determining the positional relationship between the targets and calculating the three-dimensional displacement of the seismic isolator. To improve image clarity, this paper introduces video super-resolution processing technology. Real-time hybrid experiment results show that, despite the increased measurement error due to camera movement, the displacement measurement accuracy of this method is still controlled within 2.5mm, demonstrating its effectiveness under the impact of earthquakes.

Keywords: Seismic isolation bearing; Monocular vision; PnP algorithm; Video super-resolution; 3D displacement

1 Introduction

Seismic isolators, subject to environmental loads, seismic effects, material aging, and human factors, are prone to deformation and damage. As vertical load-bearing components and vibration energy dissipation devices, they are often the most vulnerable key components in a structure. Therefore, rapid and direct measurement of the displacement of seismic isolators during earthquakes is crucial for assessing the performance of existing seismic isolation structures and bearings.

Traditional methods of displacement measurement for seismic isolators^[1,2], such as accelerometers and displacement sensors, are typically contact-based, making monitoring systems expensive and maintenance complex, thus limiting their application to a few engineering projects. In recent years, vision-based measurement methods have become popular due to their non-contact nature, multi-point monitoring capabilities, low cost, and high precision. However, these methods generally require fixed cameras, which may also move during an earthquake. Although GPS or INS systems can be used

to track camera position, GPS signals are weak indoors, and high-precision INS systems are costly, negating the cost advantage. Hence, literature^[3] introduced a computer vision-based method for measuring the displacement of seismic isolators. This method involves installing targets on the upper and lower connecting plates of the isolator and using two orthogonally positioned cameras to track the target displacements, thereby measuring horizontal and vertical displacements. Despite its accuracy, the need for two cameras increases costs and makes installation more inconvenient, limiting practical application.

Recent studies have employed a single camera to measure the three-dimensional displacement of a structure. For example, Chang et al.^[4] tracked multiple points on a planar pattern to measure the three-dimensional translation and rotation of a target attached to a structure, while Lin et al.^[5] used a single camera to capture circular and sinusoidal stripe patterns, analyzing these patterns' changes to measure three-dimensional vibrations. Although these methods are precise, they require the camera to be stationary and have high resolution. Addressing the monitoring needs of seismic isolators, this paper proposes a new method that utilizes a monocular camera and targets located above and below the isolator to track its displacement, thereby measuring the tri-directional displacement of the seismic isolator during an earthquake. Compared to dual-camera systems, this method only uses one camera, reducing hardware costs and simplifying installation. Unlike other monocular vision methods, our approach can still perform measurements even when both the camera and the seismic isolator are in motion.

2 Monocular Vision Measurement Method for Displacement of Base Isolators

2.1 Camera Setup and Target Design

To measure the displacement of seismic isolators, the camera is fixed on the beam top or ground near the isolator, and both the camera and the target move during an earthquake. To ensure measurement accuracy, artificial targets are set up along the horizontal direction of the structure, designed as circular patterns with a 3x3 chessboard grid. Due to the typically low indoor lighting conditions at the seismic isolation layer, green is chosen for the circle's color to minimize interference. The diameter of the circle is 60mm, and each square of the chessboard grid measures 12.5mm x 12.5mm. As shown in Figure 1:

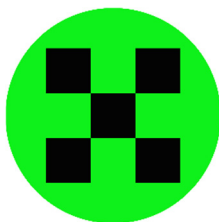


Fig. 1. Target Pattern

The targets are placed at the upper and lower piers of the seismic isolator. Thus, the deformation of the isolator during an earthquake can be represented by the relative displacement of the upper and lower targets, with the target plane perpendicular to the camera's optical axis. The arrangement of the camera and target points for measuring the displacement of the seismic isolator is shown in Figure 2:

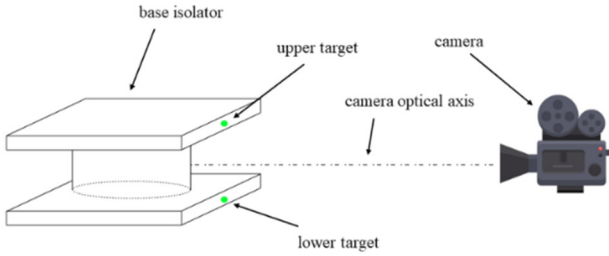


Fig. 2. Schematic of the Camera and Target Position Relationship

2.2 Target Recognition and Positioning

Firstly, the camera is calibrated to obtain the intrinsic matrix, and then used to capture video of the seismic isolator, which is subsequently split into static images for preprocessing. This preprocessing includes the application of Gaussian and median filtering to reduce noise. Subsequently, each image undergoes HSV color transformation to extract the green of the target, creating a binary mask. By analyzing the area and aspect ratio of connected components, the positions of the upper and lower targets are identified. Sub-pixel level Shi-Tomasi algorithm is then used to precisely identify the chessboard corners, selecting the top-left corner as the coordinate origin to locate the 16 corners of the chessboard grid. Due to insufficient clarity of the features at the outermost four corners, these four corner points are removed, retaining 12 corner points. The detection scenario is illustrated in Figure 3.

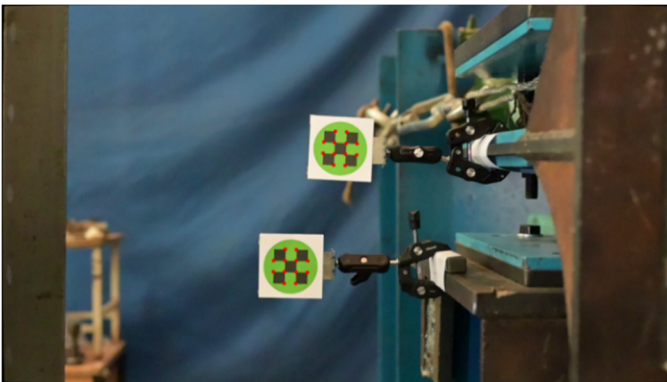


Fig. 3. Corner Detection Algorithm

2.3 Calculation of Three-Dimensional Displacement of Base Isolators

When using the camera imaging model, any three-dimensional scene can be transformed into a two-dimensional image through perspective projection. This process typically involves a linear camera model, as shown in Figure 4. In this model, $o_c x_c y_c z_c$ represents the camera's coordinate system, $o_1 xy$ is the image coordinate system of the image plane, $o_2 uv$ is the pixel coordinate system of the image plane, and $O_w X_w Y_w Z_w$ denotes the coordinate system of the observed world.

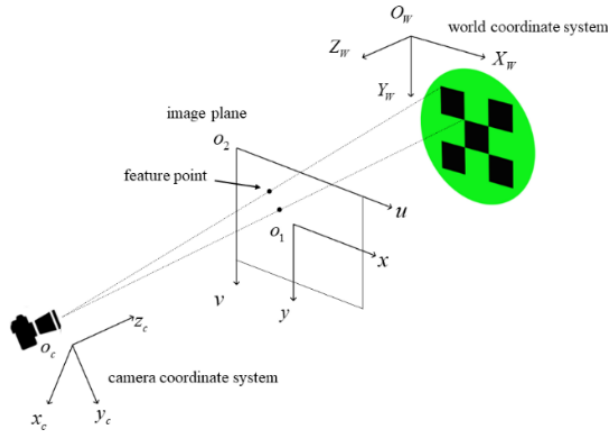


Fig. 4. Linear Camera Model

In the linear camera model, there is a mapping relationship between a point in the world coordinate system and a point in the pixel coordinate system. Specifically, when the coordinates of this point in the world coordinate system are (X, Y, Z) , correspondingly, this point can be mapped to the corresponding point (u, v) in the pixel coordinate system through perspective projection. This projection mapping relationship can be formalized using homogeneous coordinates and matrix transformations:

$$z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \mathbf{KM} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (1)$$

In the formula, f_x, f_y, u_0, v_0 are internal parameters of the camera, which can be obtained through camera calibration. R is the rotation matrix, a 3×3 orthogonal unit matrix. t is the three-dimensional translation vector from the origin of the world coordinate system to the origin of the camera coordinate system. K is the camera's intrinsic parameter matrix, and M is the camera's extrinsic parameter matrix.

Having mastered the internal parameters of the camera and acquired the three-dimensional world coordinates of sufficient spatial points along with their corresponding two-dimensional pixel coordinates on the image, the linear camera model can be used to estimate the rotation matrix and translation vector. This process essentially involves solving the Perspective-n-Point (PnP^[6]) problem.

In a monocular vision system, by setting a reference plane as the world coordinate system base point and using the PnP algorithm, it is possible to calculate the three-dimensional displacement of an object. In measuring the displacement of seismic isolators, the earthquake causes both the camera and the upper and lower plates of the isolator to move relative to the ground. The measurement focuses on the relative displacement between these plates. Therefore, it is necessary to establish an independent world coordinate system for each target, and by measuring the difference in distance between the camera and the upper and lower plates, the three-dimensional displacement of the seismic isolator can be precisely obtained.

From the solution of the PnP algorithm, the rotation matrices and translation vectors for the upper and lower targets relative to the camera are obtained. The pose of the upper target relative to the camera is represented as a transformation matrix:

$$M_{t_cam} = \begin{bmatrix} R_t & t_t \\ 0^T & 1 \end{bmatrix} \quad (2)$$

Similarly, the pose of the lower target relative to the camera can be represented as a transformation matrix:

$$M_{b_cam} = \begin{bmatrix} R_b & t_b \\ 0^T & 1 \end{bmatrix} \quad (3)$$

To obtain the pose of the upper target within the coordinate system of the lower target, it is first necessary to find the pose of the camera in the coordinate system of the lower target. This can be achieved by taking the inverse of M_{b_cam} . Additionally, since the rotation matrix is orthogonal:

$$M_{cam_b} = M_{b_cam}^{-1} = \begin{bmatrix} R_b^T & -R_b^T t_b \\ 0^T & 1 \end{bmatrix} \quad (4)$$

By combining the transformation of the upper target relative to the camera with the transformation of the camera in the coordinate system of the lower target, the pose of the upper target in the coordinate system of the lower target is obtained:

$$M_{t_b} = M_{cam_b} \cdot M_{t_cam} = \begin{bmatrix} R_b^T & -R_b^T t_b \\ 0^T & 1 \end{bmatrix} \cdot \begin{bmatrix} R_t & t_t \\ 0^T & 1 \end{bmatrix} = \begin{bmatrix} R_b^T R_t & R_b^T (t_t - t_b) \\ 0^T & 1 \end{bmatrix} \quad (5)$$

Taking point $P_t = (X_t, Y_t, Z_t)$ in the upper target coordinate system and $P_b = (X_b, Y_b, Z_b)$ in the lower target coordinate system, the coordinate relationship

between the upper and lower targets can be determined through the pose of the upper target in the coordinate system of the lower target:

$$\begin{bmatrix} X_b \\ Y_b \\ Z_b \\ 1 \end{bmatrix} = \begin{bmatrix} R_b^T R_t & R_b^T (t_t - t_b) \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} X_t \\ Y_t \\ Z_t \\ 1 \end{bmatrix} \quad (6)$$

Both the upper and lower targets choose their coordinate system origins to be located on the targets, denoted as $P_i = (0,0,0)$. By doing so, the coordinates of the origin of the upper target's coordinate system in the lower target's coordinate system can be obtained. Where t_{ti} and t_{bi} are the components of the translation vectors t_t and t_b respectively, and R_{bij} are the elements of the rotation matrix R_b .

$$\begin{aligned} P_b &= \begin{bmatrix} R_b^T R_t & R_b^T (t_t - t_b) \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} R_b^T (t_t - t_b) \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} R_{b11}(t_{t1} - t_{b1}) + R_{b21}(t_{t2} - t_{b2}) + R_{b31}(t_{t3} - t_{b3}) \\ R_{b12}(t_{t1} - t_{b1}) + R_{b22}(t_{t2} - t_{b2}) + R_{b32}(t_{t3} - t_{b3}) \\ R_{b13}(t_{t1} - t_{b1}) + R_{b23}(t_{t2} - t_{b2}) + R_{b33}(t_{t3} - t_{b3}) \\ 1 \end{bmatrix} \end{aligned} \quad (7)$$

3 Video Super-Resolution Processing for Base Isolator Displacement Using Deep Learning Methods

3.1 Dataset Preparation

In performing visual measurement tasks, the quality of video data is a key factor in obtaining accurate measurement results. High-quality video can significantly improve the accuracy and reliability of structural displacement detection. Particularly in challenging environments like seismic isolation layers, video quality is affected by camera motion, which reduces clarity and causes motion blur, ultimately impacting the precision of measurement results. To mitigate this effect, this paper introduces video super-resolution technology to enhance the accuracy and effectiveness of the entire measurement process.

To train the BasicVSR++^[7] network model, recorded 43 video clips of seismic isolators under various earthquake motions, with the videos captured while the camera was stationary. These videos had resolutions of 1920x1080 (120 fps) and 3840x2160 (30 fps). After downsampling and adding Gaussian blur and noise, low-resolution

videos were created, as shown in Figure 5. The dataset comprises 1064 training sequences, 118 validation sequences, and 146 test sequences.

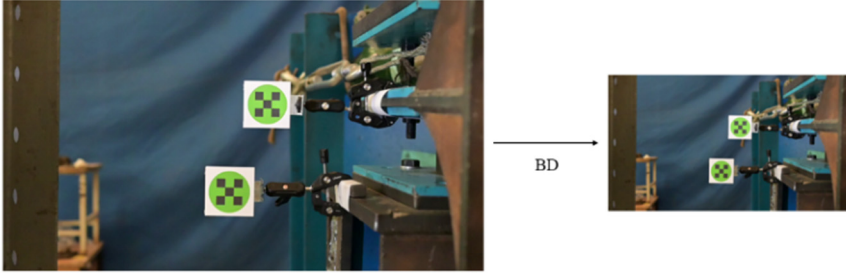


Fig. 5. Image Downsampling

3.2 Model Training

The experiment utilized the MMEediting platform based on PyTorch. The training configuration spanned 100 epochs, employing the Adam optimizer with an initial learning rate of 0.001 and beta parameters set to (0.9, 0.99). The Charbonnier loss function was selected as the loss function for training, with a loss weight of 1.0, and mean reduction was used to optimize pixel-level reconstruction quality. Balancing hardware constraints and efficiency considerations, the batch size was set to 1.

3.3 Model Evaluation and Testing

To evaluate the effectiveness of video super-resolution, this study primarily utilized two metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). PSNR assesses image quality by comparing the pixel intensity differences between the original and reconstructed images, while SSIM evaluates the structure and texture of images, more accurately reflecting the human eye's perception of detail. Test results showed that the model achieved a PSNR of 30.2146 dB and an SSIM of 0.9034, demonstrating its effectiveness in terms of pixel accuracy and structural preservation.

The video data that requires displacement measurement is processed using a trained model for video super-resolution, and then the displacement measurement is performed. Figure 6 compares the super-resolution reconstruction effects of the seismic isolator displacement detection video. The left side shows the image before reconstruction, and the right side displays the image after reconstruction, which reveals richer details and improved visual effects.

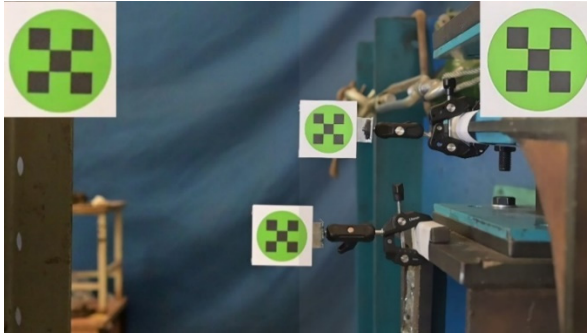


Fig. 6. Comparison of Video Before and After Super-Resolution Reconstruction

4 Experimental Verification

4.1 Experimental Setup

This study's experimental section is based on the displacement measurement of base isolators using real-time hybrid testing. The experiments were conducted on LRB200 rubber base isolators using a specific earthquake wave, and videos of the entire process were recorded to verify the effectiveness and accuracy of the proposed method. Since the equipment used in the experiment could only impose displacement in one direction, the experiment focused on the most challenging of the three-dimensional displacements—the depth direction.

In this experiment, testing control and load equipment from MTS Company were used. To simulate the performance of a camera under seismic activity, a small-scale vibration platform model WS-Z30-50 was utilized. During video capture, a Nikon Z30 camera was employed, set to a resolution of 1920×1080 and a frame rate of 120 frames per second, ensuring that the camera's optical axis was parallel to the target's normal. The experimental site is shown in Figure 7.

In the experiment, the actuator loaded the earthquake motion file, as shown in Table 1, and filming was divided into two scenarios: First, under static conditions, the camera was fixed to the ground to record the unidirectional movement of the target to analyze the displacement characteristics of the seismic isolator. Second, the camera was mounted on a vibration table to simulate the effects of an earthquake, recording the impact of seismic motion on the target and assessing performance under dynamic conditions.

Table 1. Earthquake information

Earthquake wave filename	MORGAN_SJL360.AT2
Earthquake Seismic station	Morgan Hill San Justo Dam
Magnitude	6.19
Fault distance(km)	31.88
Vs30(m/s)	543.63
Characteristic period(s)	0.38

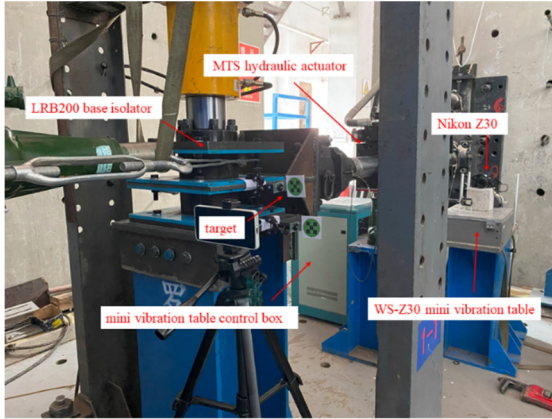


Fig. 7. Test Site Image

4.2 Experimental Results and Analysis

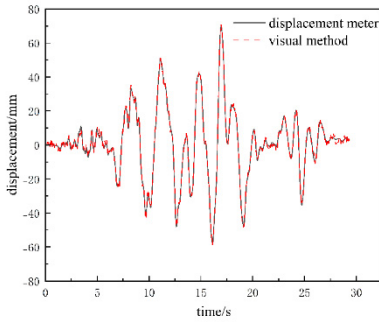


Fig. 8. Displacement time history with camera stationary

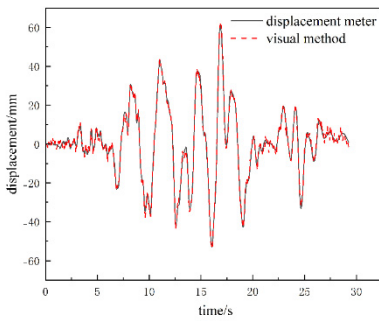


Fig. 9. Displacement time history with camera in motion

Using the described visual method, the base isolator's displacement history was obtained and compared with displacement meter readings. Figure 8 shows the stationary camera scenario, while Figure 9 shows the camera in motion. As shown in Figures 8, and 9, whether the camera is stationary or in motion, the displacement time histories obtained using the visual method proposed in this paper are essentially consistent with the results obtained by the displacement meter. The error in displacement peaks is shown in Table 2.

Table 2. Comparison of Horizontal Displacement Peaks of Base Isolators Obtained by Different Measurement Methods

Test Condition	Displacement Meter		Visual Method	
	Displacement Peak (mm)	Displacement Peak (mm)	Error Value (mm)	Error Rate (%)
Camera Stationary	69.716	71.160	1.444	2.071
Camera in Motion	61.009	63.114	2.104	3.449

As shown in Table 1, compared with the results measured by the displacement meter, the measurement errors using the visual method proposed in this paper are all less than 4%. The maximum absolute error in horizontal displacement peaks is 2.104mm, which fully meets the accuracy requirements for measuring horizontal displacement of base isolators. This demonstrates that the proposed visual method can effectively perform horizontal deformation detection of base isolators under various conditions.

5 Conclusion

This paper proposes a method for three-dimensional displacement detection of base isolators, which combines image feature recognition technology and the PnP algorithm to achieve automatic detection and positioning of target markers. The introduction of video super-resolution technology effectively improves the quality of the video, validating its effectiveness in testing the depth direction displacement of base isolators. Experimental results demonstrate that the depth direction displacement time histories obtained using this method are nearly identical to those obtained by displacement meters. Compared with the results obtained by displacement meters, the absolute error in horizontal displacement peaks is 2.104 mm, with an error ratio of less than 4%, indicating that the method can accurately measure the displacement of base isolators, meeting the accuracy requirements and demonstrating practical utility.

References

1. Zhu W., Zhang J. (2014) Real-time Monitoring Research of the Seismic Isolation System at the Guangdong Science Center Based on Intelligent Sensing Technology. *Journal of Civil Engineering*, 47 (05): 40-45. <https://doi.org/10.15951/j.tmgcxb.2014.05.013>.

2. Li W., Zheng W., Du Y., et al. (2016) Design and Implementation of Health Monitoring System for Base-Isolated Structures (I): System Design. *Journal of Earthquake Engineering*, 38 (01): 94-102. <https://www.cqvip.com/qk/92873a/201601/668602689.html>
3. Dang Y., He Y. (2023) Dynamic Displacement Measurement Method for Seismic Isolators Based on Computer Vision and Deep Learning. *Journal of Vibration and Shock*, 42 (06): 90-97+165. <https://doi.org/10.13465/j.cnki.jvs.2023.06.011>.
4. Chang C. C., Xiao X. H. (2010) Three-dimensional structural translation and rotation measurement using monocular videogrammetry. *Journal of Engineering Mechanics*, 136 (7): 840-848. [https://doi.org/10.1061/\(ASCE\)EM.1943-7889.0000127](https://doi.org/10.1061/(ASCE)EM.1943-7889.0000127).
5. Lin Y., Zhong S., Zhong J., et al. (2021) Three-dimensional Vibration Measurement Based on Combined Feature Patterns in Monocular Vision. *Laser & Optoelectronics Progress*, 58 (22): 336-345. <https://doi.org/10.3788/LOP202158.2212002>.
6. Terzakis G., Lourakis M. (2020) A consistently fast and globally optimal solution to the perspective-n-point problem. In: *Computer Vision—ECCV 2020: 16th European Conference*. Glasgow, UK. pp. 478-494. https://doi.org/10.1007/978-3-030-58452-8_28.
7. Chan K. C. K., Zhou S., Xu X., et al. (2022) BasicVSR++: Improving Video Super-Resolution with Enhanced Propagation and Alignment. In: *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, LA, USA. pp. 5962-5971. <https://doi.org/10.1109/CVPR52688.2022.00588>.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

