Article

Multi-Cell Displacement Measurement During the Assembly of Automotive Power Batteries Based on Machine Vision

Yueda Xu [1], Yanfeng Xing [1,*], Hongbo Zhao [2], Yufang Lin [3], Lijia Ren [3] and Zhihan Zhou [4]

1 School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201600, China; m310122432@sues.edu.cn
2 Pan Asia Technical Automotive Center Co., Ltd., Shanghai 201201, China; hongbo_zhao@patac.com.cn
3 SAIC GM Power Technology (Shanghai) Co., Ltd., Shanghai 201206, China; yufang_lin@saic-gm.com (Y.L.); lijia_ren@saic-gm.com (L.R.)
4 Avatr Technology (Chongqing) Co., Ltd., Chongqing 401120, China; zhihan.zhou@avatr.com
* Correspondence: xyf2001721@sues.edu.cn

Abstract: The positioning of lithium battery tabs in electric vehicles is a crucial aspect of the power battery assembly process. During the pre-tightening process of the lithium battery stack assembly, cells and foams undergo different deformations, leading to varying displacements of cells at different levels. Consequently, determining tab positions poses numerous challenges during the pre-tightening process of the stack assembly. To address these challenges, this paper proposes a method for detecting feature points and calculating the displacement of lithium battery stack tabs based on the MicKey method. This research focuses on the cell tab, utilizing the hue, saturation, and value (HSV) color space for image segmentation to adaptively extract the cell tab region and further obtain the ROI of the cell tab. In order to enhance the accuracy of tab displacement calculation, a novel method for feature point detection and displacement calculation of lithium battery stacks based on the MicKey (Metric Keypoints) method is introduced. MicKey can predict the coordinates of corresponding keypoints in the 3D camera space through keypoint matching based on neural networks, and it can acquire feature point pairs of the subject to be measured through its unique depth reduction characteristics. Results demonstrate that the average displacement error and root mean square error of this method are 0.03 mm and 0.04 mm, respectively. Compared to other feature matching algorithms, this method can more consistently and accurately detect feature points and calculate displacements, meeting the positioning accuracy requirements for the stack pole ear in the actual assembly process. It provides a theoretical foundation for subsequent procedures.

Keywords: lithium battery tabs; machine vision; displacement measurement; HSV

1. Introduction

As a relatively mature energy storage technology, the performance and lifespan of lithium batteries are influenced by various factors. To address these challenges, it is essential to establish a comprehensive battery health management system. One significant feature of such a health management system, compared to traditional Battery Management System (BMS), is the development of a dimension–preload force–expansion force model. This model utilizes the stress and volumetric energy density indicators of individual battery cells to optimize the arrangement of cells within the battery pack, thereby enhancing its durability. A critical issue to be addressed within the size–preload–expansion force model is the precise control of individual cell displacement and accurate positioning

during the preloading phase of a stacked assembly. Additionally, during the assembly process of the battery pack, individual cells are connected using various techniques such as resistance welding, laser welding, ultrasonic welding, and mechanical connections, effectively integrating the components of the battery pack [1]. Among these, ultrasonic metal welding, resistance spot welding, and pulse tungsten inert gas (TIG) spot welding are the three preferred welding methods for connecting batteries through tabs or busbars. This preference is due to the high efficiency of spot welding in the small-scale connection process of battery stack components [2]. Due to the stacking of multiple layers of battery cells separated by foam pads of varying specifications, the application of preload can result in varying degrees of displacement and deformation among the different layers of battery cells and foam pads. These slight displacements may create uneven gaps between the battery cells, further complicating the precise positioning required for subsequent tab welding. In some cases, this may even lead to poor contact between layers or localized internal stress concentration, ultimately degrading battery performance and introducing safety risks [3–5]. Therefore, accurately detecting and measuring the displacement of individual battery cells during the assembly process is crucial. By implementing real-time monitoring and control of battery cell displacement, it becomes possible to dynamically adjust assembly parameters and processes, ensuring the stability and uniformity of the battery cells. This approach not only enhances the performance and lifespan of lithium battery packs but also reduces the overall costs and safety risks associated with energy storage systems.

Existing displacement detection methods are primarily categorized into contact and non-contact measurements [6]. Contact measurement techniques have been widely utilized in various industrial inspections, such as structural displacement monitoring in buildings [7]. However, these methods face significant limitations. Firstly, installing contact measurement sensors on flexible surfaces is often challenging. Secondly, contact measurement requires direct physical interaction with the object being measured, which can introduce additional deformations, leading to inaccuracies. As a result, it is difficult to apply contact measurement techniques effectively in the cell assembly process. Non-contact measurement methods are better suited to the conditions of cell assembly. Common non-contact approaches include laser measurement [8], ultrasonic measurement [9], and visual measurement [10]. These methods have now reached a mature level of development and are applicable across various industries in numerous scenarios. In particular, industrial laser displacement measuring instruments are widely used for displacement measurement, offering accuracy that can easily reach 0.01 mm depending on the specific model, along with a range of measurement distances available. However, a significant drawback of these instruments is their high cost; for monitoring multi-story structures, multiple units must be purchased, necessitating ample space for installation and calibration, which imposes certain limitations. As a result, these instruments are not particularly suitable for the scenarios explored in this study.

In the process of battery pack assembly, which involves complex multi-layer deformations and various structural changes, visual measurement methods present a more practical alternative compared to other sensors that entail high costs and cumbersome installation processes. Visual measurement requires merely a stationary image acquisition device, with displacement being determined through the analysis of captured images that illustrate the relative positions of cell monoliths across different layers. Moreover, the visual measurement approach offers a considerable advantage over contact displacement sensors: each pixel in the captured image can be treated as a sensor, thus facilitating high-resolution analysis. Based on this principle, this paper proposes a method for multi-layer displacement detection of battery monomers utilizing a standard camera, which proves to be both cost-effective and suitable for engineering applications.

Non-contact measurement based on machine vision has been widely applied in various engineering fields, particularly in the study of displacement in bridges and buildings. Piotr Olaszek [11] proposed an edge crossing detection algorithm to identify special markers placed on the target structure, enabling displacement measurement for various types of bridges under different conditions. The feasibility of this approach was validated through experiments and comparisons with displacement meters. White K.R. et al. [12] from New Mexico State University utilized the close-range photogrammetry method (DCRTP) to perform both laboratory and real-world bridge experiments. By applying elastic beam theory, they verified the accuracy of the measurement data, and further demonstrated the practicality of the method through bridge static load experiments. Yu Shanshan et al. [13] developed a fast feature tracking software for multi-objective deformation measurement, which significantly improved feature extraction speed by reducing the requirements for scale and rotation invariance. Additionally, the traditional RANSAC algorithm was optimized by pre-filtering matched point pairs and quickly discarding unreasonable parameter models, improving both the accuracy and efficiency of the SURF-BRISK-based feature tracking method. Xu et al. [14] employed a feature matching method for random five-marker detection on the external features of a pedestrian bridge. This enabled the measurement of dynamic displacement and vibration variations under different loading conditions. Dongming Feng et al. [15] used contour detection and feature matching techniques to measure the displacement of a cable-stayed bridge. Their method also facilitated the measurement of cable vibrations and bridge deck displacements. In subsequent research, Dongming Feng [16] developed a new visual sensor system for remote displacement measurement of structures. This system integrated a template matching algorithm with the OCM (Oriented Corner Matching) algorithm and the UCC (Unbiased Correlation Coefficient) algorithm to effectively detect and track feature points on structural surfaces. The improved upsampling factor achieved better sub-pixel resolution, enhancing the system's robustness even in harsh environments.

Based on the aforementioned work, it is evident that research on displacement measurement utilizing visual methods has become quite widespread. However, several challenges persist in practical applications. Firstly, while there is a variety of visual measurement methods available, including traditional recognition algorithms based on image feature point matching and emerging machine learning technologies, the horizontal application of these methods remains limited. Furthermore, there is often a lack of practical scenario implementations and quantitative analyses in existing review articles, which restricts their guidance on actual engineering production processes. Secondly, machine vision-based displacement measurement techniques are predominantly used for vibration and displacement monitoring of large structures, such as bridges and buildings, with a strong emphasis on real-time and long-term monitoring. In contrast, during the pre-tensioning process of electrical stacks, particular attention must be paid to the accuracy of displacement detection at specific moments. In contrast to the long-distance macro-structural monitoring of bridges and buildings, the application of machine vision for monitoring displacement changes in multi-cell configurations during the pre-tensioning process—a scenario that involves close-range, multi-story structural displacement detection—has received relatively little attention. Furthermore, the majority of industry research focuses on cell expansion displacement during the charge and discharge cycles following pre-tensioning [17,18], while insufficient emphasis is placed on the core displacement occurring during the pre-tensioning phase, prior to the mechanical connection of the core to other components. In the context of automated production processes, the accurate positioning of cells and the measurement of their displacement are undoubtedly crucial for enhancing the overall quality of the final battery pack product.

Therefore, analogous to the deformation monitoring of floor structures and bridge structures, a similar approach can be adopted for the process of multi-cell stacking pre-tensioning. To address the challenges of detecting the displacement of multi-layer cells during the assembly of lithium batteries under complex working conditions, this paper proposes a novel method for cell displacement detection based on the MicKey method, which eliminates the need for special markers. First, the region of interest (ROI) of the tab is identified using an adaptive cell boundary segmentation approach based on the HSV color space. Next, the MicKey method's neural network keypoint matching process is employed to predict the coordinates of corresponding keypoints in the 3D camera space, reconstruct the 3D target object from 2D images, and extract and match feature point pairs. Finally, the Z-Score method is applied to filter out outlier mismatched points, enabling the calculation of accurate pixel displacements. These pixel displacements are then used to determine the precise displacement of tabs at each cell level during the assembly process. This method is integrated into a comprehensive health management system for lithium battery production, and its performance has been evaluated through a series of laboratory and field tests. The results demonstrate its effectiveness and practicality in accurately detecting multi-layer cell displacements during assembly.

The structure of this paper is as follows: Section 2 provides a detailed description of the overall process of the proposed method. It introduces the HSV image segmentation technique for extracting the target region (ROI), along with the MicKey feature matching process and the scaling factor estimation method. In Section 3, the proposed method is validated through its application in the actual stack assembly process, and its advantages over other feature matching methods are analyzed. Finally, Section 4 presents the conclusions and discusses potential directions for future research.

2. Materials and Methods

To address the challenge of conveniently measuring lithium battery stacks during the pre-tightening process, this paper proposes a method for detecting cell displacement during lithium battery assembly based on the MicKey method. First, an HSV-based approach is applied to automatically extract the ROI of the cell tabs, effectively reducing computational complexity. Next, the neural network keypoint matching process of the MicKey method is employed to track the target feature points, and the Z-Score method is utilized to filter out outlier feature points, ensuring accurate pixel displacement of the feature points. Finally, the actual displacement is calculated by converting the pixel displacement using the scaling factor formula, enabling precise measurement of the cell tab displacement. The flowchart of the proposed method is shown in Figure 1.
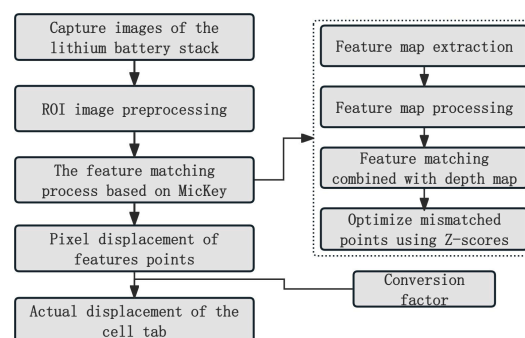

Figure 1. Flowchart of displacement detection based on machine vision.

It is important to note that under actual operating conditions, the battery stack experiences deformations beyond just those aligned with the direction of pre-tensioning force.

For instance, the tabs (or terminals) may exhibit small displacements in the z-direction, as illustrated in Figure 2. These displacements arise partly from the compressive forces exerted by the pre-tensioning, leading to deformation of the cell casing, and partly because the cell casing is not an ideally flat surface. Additionally, the pre-tensioning direction is not entirely perpendicular to the stacking direction of the cells, which can result in lateral slipping of certain cells within the battery stack in the z-direction. However, given that the predominant deformation still occurs along the direction of the pre-tensioning force, this paper focuses primarily on the displacement of cells in the pre-tensioning direction during the battery pack assembly phase, specifically the displacements along the y-direction, as depicted in the subsequent figure.
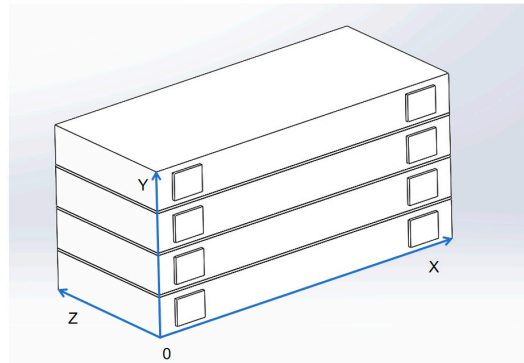

Figure 2. Schematic representation of the deformation direction of the cell.

2.1. ROI Selection Based on HSV

The assembly environment of the battery cell includes various elements beyond the battery cell itself, such as fixtures, foam cushions, presses, and other background components, as illustrated in Figure 3.

Figure 3. Lithium battery stack assembly environment.

These elements can interfere with the displacement measurement of the battery cell and negatively impact the accuracy of the measurement results. Power batteries are typically covered with a blue film, where the blue color serves as the contour boundary of a single battery cell. This color is significantly distinct from the background, making it easier to identify the cell. In this paper, the HSV color space is utilized to differentiate the body of the battery cell, covered with a blue film, from its background. Compared to the RGB color space, the HSV color space is less sensitive to lighting variations, as changes in saturation and brightness do not affect the hue component [19]. Thus, the HSV color space is well suited for separating the blue outline of the battery cell from the background. By extracting the HSV histograms of the blue film covering the lithium battery pack under various lighting conditions, the interval distribution of the target pixels in the histogram is analyzed, and the values of H, S, and V are determined, as shown in Equation (1).

Specifically, when the hue (H) is within the range [100–180], the saturation (S) is within [180–255], and the brightness (V) remains at its default value, the segmentation of the target becomes more precise.

$$minBlue = [H\_min, S\_min, V\_min]$$
$$maxBlue = [H\_max, S\_max, V\_max]$$

(1)

The contour extraction map resulting from the segmentation of the main body of the battery cell using the HSV color space is shown in Figure 4a. This map is further processed through binarization, and the contours are made continuous to a certain extent using morphological processing. The final contour extraction map of the battery stack is obtained, as shown in Figure 4b below.
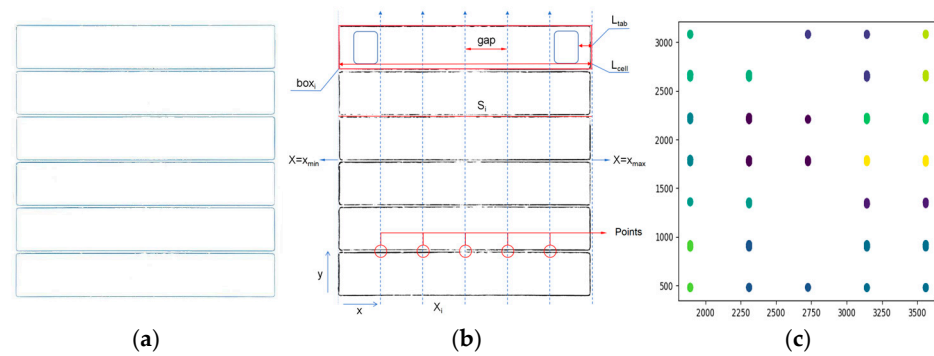

(**a**)　　　　　(**b**)　　　　　(**c**)

Figure 4. Flow chart of cell contour extraction: (**a**) rough extraction of the main contour of the battery cell; (**b**) morphological post-processing of contours and cell boundary detection methods. (**c**) Boundary point detection result.

At this stage, although the contour has been processed, it remains intermittent, and traditional edge detection methods, such as Canny, fail to produce satisfactory results. To address this, the method illustrated in Figure 4b is applied. The x-value range of the contour $[x_{min}, x_{max}]$ is obtained by scanning each column to determine the left and right boundaries of the rectangular core. Next, the contour of the battery cell is sampled at intervals by defining detection lines X_n, as shown in Equation (2). The value of i can be adjusted based on specific requirements; in this study, i is set to 5. The sampling points captured by X_n are denoted as points. The sampled data are then classified based on the ordinate values to generate a scatter plot of the upper and lower boundaries of the battery cell. The scatter plot is grouped according to the distribution of the ordinate values, and the mean value of each group is calculated to determine the dividing lines L_i between each cell level, as shown in Equation (3).

$$\begin{cases} X_n = x_{min} + gap \times n \quad n = (0, 1, \cdots i) \\ gap = \frac{x_{min} + x_{max}}{i+1} \end{cases}$$

(2)

$$L_i = \frac{\sum_{j=1}^{N} y_j}{N}$$

(3)

where X_n is the current rectangular contour detection line, n is the number of gaps, i is a preset value used to adjust the density of the detection lines, and N is the total number of detected points.

Using X_{min}, X_{max}, and L_i, the four boundaries of the simplified rectangular outline corresponding to each cell layer, or box, can be determined. The values of L_{tab} and L_{cell} are fixed for each batch of cells, allowing the position of the tab within the simplified outline to

be calculated based on these two parameters. Consequently, the area of the cell tab can be identified and defined as the ROI area, as illustrated in Figure 5.

Figure 5. Extracted area of the lug.

2.2. Feature Matching Process

This study employs the MicKey algorithm to establish precise 3D-3D correspondences in camera space from 2D images through descriptor matching, addressing two critical challenges in feature detection and matching. Metric Keypoints (MicKey) [20] directly predict keypoint locations in camera space, forming metric correspondences without requiring explicit depth measurements or global structural information. These correspondences enable the recovery of relative pose between two views through differentiable pose optimization, allowing end-to-end training of MicKey using only image pairs and their ground truth relative poses for supervision.

Unlike traditional methods that rely on depth measurements or structure-from-motion (SfM) reconstruction, MicKey focuses on regions with reliable features, inherently learning depth information only where it is meaningful. This weakly supervised approach eliminates the need for additional overlap information or global structure reconstruction, significantly enhancing accessibility and practicality. For new domains, MicKey requires only pose data for training, bypassing the need for extensive domain-specific preprocessing or calibration.

MicKey adopts a multi-headed network architecture with shared encoders [21–23], as illustrated in Figure 6. The shared encoder leverages a pre-trained DINOv2 [24] network to extract features. The image is segmented into 14×14 blocks, each represented by a feature vector, forming a feature map $F \in \mathbb{R}^{1024 \times w \times h}$, where $w = W/14$ and $h = H/14$. The multi-head design facilitates parallel prediction of the following:

- U(2D offset): Calculates the 2D location of keypoints relative to block centers.
- C(Confidence): Indicates the reliability of detected keypoints.
- Z(Depth): Estimates depth for each keypoint.
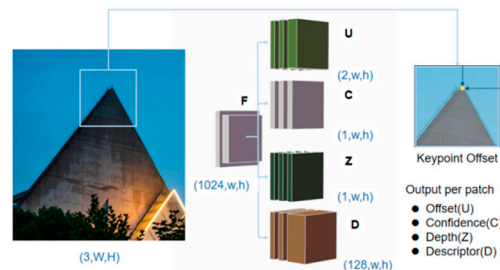- D(Descriptor vector): Encodes unique feature descriptors for matching.

Figure 6. MicKey feature extraction workflow.

This configuration directly outputs 3D Metric Keypoints, enabling descriptor-based matching between images without requiring explicit depth supervision.

Once the keypoint matching data is obtained from an image pair, the Z-Score method is utilized to eliminate outliers. The Z-Score for each data point is calculated using the

displacement values $\{d_i\}$, following Equation (4), which incorporates the mean and standard deviation.

$$z_i = \frac{d_i - \mu}{\sigma} \tag{4}$$

Outliers with $|z_i| > \tau$ (predefined threshold) are excluded, ensuring robust matching. This statistical filtering enhances the reliability of correspondences, especially under varying conditions or noise.

To further validate the reliability of the proposed method, a manual evaluation was conducted on a subset of 100 randomly selected image pairs from the dataset. Three independent experts were tasked with manually annotating the keypoint correspondences and the resulting displacements. The manually annotated displacements were averaged to serve as a reference standard. Comparing the results of the MicKey algorithm with these manually annotated ground truths, the method achieved a mean absolute error of 0.031 mm and an RMSE of 0.041 mm, consistent with the laser displacement sensor measurements. These findings demonstrate that the MicKey algorithm aligns closely with human judgment and is robust against variations in conditions, further reinforcing its reliability.

In addition, the proposed method was validated using a proprietary dataset to measure displacements during lithium battery cell assembly. Keypoint matches obtained by MicKey were compared against reference measurements from a laser displacement sensor, regarded as ground truth. For each image pair, the pixel displacement d_{pixel} calculated by MicKey was converted to real-world displacement using a calibration factor in Equation (5):

$$d_{\text{real}} = CF \cdot d_{\text{pixel}}, \quad CF = \frac{d_{\text{world}}}{d_{\text{icmos}} \cdot \frac{d_{\text{pixel}}}{f}} \tag{5}$$

where d_{world} is the actual physical dimension, d_{icmos} is the sensor's pixel size, and f is the focal length.

MicKey demonstrated a mean error of 0.029 mm and an RMSE of 0.039 mm across 625 image pairs, outperforming traditional algorithms such as SURF-FLANN and ORB. Compared to traditional methods, MicKey reduced the measurement error by 52%, achieving superior accuracy and stability. Detailed results are presented in Section 3.

The MicKey algorithm introduces a weakly supervised framework capable of achieving high precision in feature detection and matching, making it highly suitable for complex industrial applications such as multi-layer displacement detection in lithium battery cell assembly. By combining advanced feature extraction via pre-trained DINOv2, robust outlier filtering using Z-Score, and differentiable pose optimization, MicKey delivers reliable results under challenging conditions. This innovative approach addresses key limitations of existing methods, particularly the reliance on explicit depth measurements or global structure reconstruction, providing a scalable solution for industrial and research applications.

2.3. Determination of the Conversion Factor

In single-camera measurement, displacement information of the structure in both the horizontal and vertical directions can typically be obtained. Since the stack structure primarily experiences displacement along the horizontal and vertical planes during the compression of the cell structure, this study focuses on measuring the vertical pixel displacement (Y_i) in the direction of the applied pressure on the cell, as illustrated in Equation (6).

$$Y_i = \frac{\sum_{j=1}^{N} \left(y_i^j - y_j^1 \right)}{N} \tag{6}$$

where y_i^j and y_j^1 and are the y-coordinate differences in the jth point pair in the ith and first image sequences, respectively, and n is the number of correctly matched points in the two images.

In structural displacement measurement, it is essential to establish a conversion ratio between pixel displacement and real-world coordinate displacement, as illustrated in Figure 7. The red lines represent the displacement of the object as well as the corresponding displacement on the image plane, while the blue line denotes the optical axis of the lens. After the field equipment is calibrated using a spirit level, the camera's imaging plane is aligned to be parallel to the surface of the object being measured, ensuring that the angle between the camera plane and the target surface is approximately $0°$. The conversion factor (CF) is then determined by calculating the ratio of the known physical dimensions of the target surface to its corresponding pixel dimensions, as expressed in Equation (5).
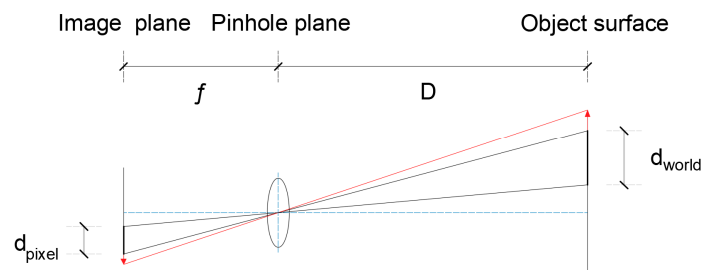

Figure 7. Conversion factor determination.

3. Results

3.1. Experimental Platform Construction and Experimental Plan Design

To validate the accuracy of the measurement method proposed in this paper, an experimental platform was constructed based on the actual working conditions of battery cells assembled into a stack, as shown in Figure 8. The platform utilizes a NIKON Z30 camera (Nikon, Tokyo, Japan) paired with a 50 mm Yongnuo lens (Yongnuo, Shenzhen, China), capable of capturing images at a resolution of 5568×3721. A multifunctional press is employed as the pressure loading device. This press is capable of applying precise pressure loads to the stack in both manual and automatic modes, with accurate control of the pressing speed and the maximum applied pressure. Additionally, the control panel continuously displays the relative position of the upper end plate and the current pressure in real time. Furthermore, this experiment employs a laser displacement sensor with a repeatability precision of 10 μm to measure the displacements of cells at different layers, using these measurements as standard reference values for displacement.

Figure 8. Experimental platform.

The experimental procedure involves operating the multifunctional press under controlled conditions. The downward travel distance and speed are preset on the main control system, and uniform pressure is applied to the stack in accordance with actual assembly

conditions. For this experiment, the downward speed is set to 0.05 mm/s. The displacement of the upper end plate relative to the target cell level, as detected by the upper end plate laser displacement detector, is recorded starting from the moment the pressure indicator on the main control panel reaches 2000 N. Simultaneously, the relationship between the applied pressure on the stack and the corresponding displacement is documented. The camera captures an image of the stack at one-second intervals throughout the process. Additionally, a laser displacement sensor measures the distance from the upper plate of the press to the marking point on the target cell. This measurement provides the precise displacement of the cell, which is used as a reference value for validating the results.

In this study, the SURF-FLANN [25] and ORB feature matching algorithms with adaptive thresholds [26] are used as comparison methods. To demonstrate that the proposed method can accurately detect displacement changes in multiple cells within a stack, experiments were conducted to measure the displacement of cells in a four-cell stack and an eight-cell stack. These experiments were designed to validate the method's effectiveness in detecting displacement across different numbers of cells, simulating conditions encountered in actual production processes. Furthermore, to assess the accuracy and stability of the proposed method, an error analysis was performed using the root mean square error (RMSE), as defined in Equation (7).

$$\text{RMSE} = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(x - \hat{x})^2} \tag{7}$$

where k is the total number of measured data, x is the calculated value of the visual method, and $\hat{x}$ is the measured reference value of the laser displacement sensor.

3.2. Multi-Level Cell Displacement Detection Experiment

To evaluate the practicality of the proposed algorithm, initial experiments were conducted using a smaller set of four-layer cells. After automatically extracting the region of interest (ROI) of the tab, the ROI was processed using the MicKey matching algorithm. The results of the feature detection and matching stage, compared with the two other methods, are shown in Figure 9. As illustrated in Figure 9a, the SURF-FLANN method detects a large number of feature pairs. However, a significant proportion of these are irrelevant points located in low-feature areas outside the tab region, leading to increased noise. In contrast, the ORB method with adaptive thresholding, as shown in Figure 9b, generates feature point pairs that are primarily located in key areas. However, the total number of detected pairs is relatively small. By comparison, and as demonstrated in the MicKey depth estimation restoration maps in Figure 9c,d, the feature point pairs extracted by the proposed method are more densely concentrated in the tab region, with significantly fewer mismatched points. This highlights the superior precision and reliability of the proposed algorithm in feature detection and matching.

Next, the feature points extracted using the three methods were used to compute the variation in the pixel Y value along the force direction. Outliers were eliminated using the Z-Score method to obtain the optimal set of matching points. The average value of these points was then calculated, yielding the results shown in Figure 10. The red points in the figure represent the filtered data, while the blue points correspond to the original raw data. These results further demonstrate the accuracy and stability of the proposed method.
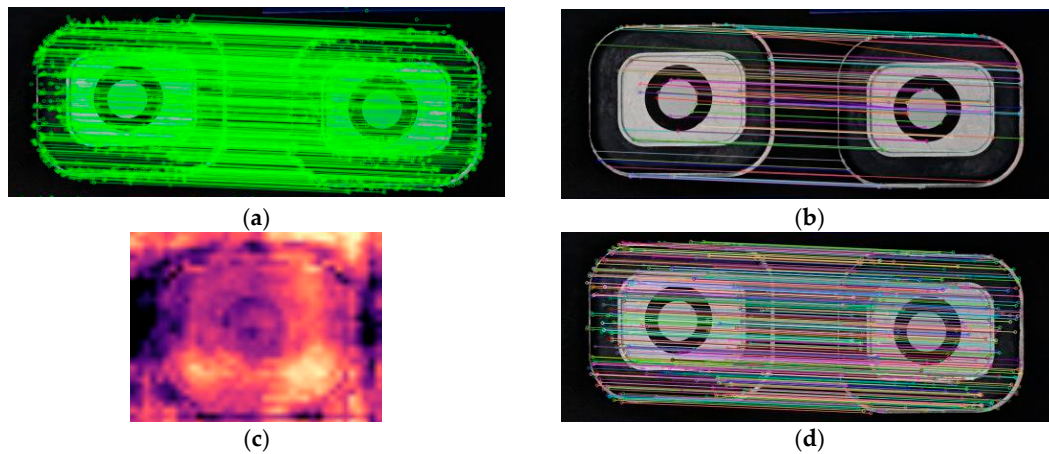
Figure 9. Detection results of different methods: (**a**) SURF-FLANN; (**b**) ORB with adaptive thresholds; (**c**) deep restoration of images by our method; (**d**) ours.
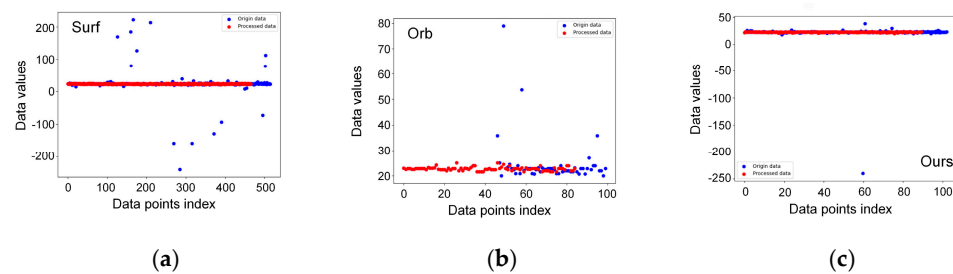

Figure 10. Results after outlier removal using Z−Score: (**a**) SURF−FLANN; (**b**) ORB with adaptive thresholds; (**c**) ours.

Using the batch number stamped on the battery as a physical marker, combined with the physical dimensions of the battery and the parameters outlined in Equation (5), the CF (conversion factor) of the current system is calculated to be 0.069 mm/pixel.

To evaluate whether the accuracy of the proposed algorithm meets the required standards, the displacement detection of four cell tabs under a 2000 N load was used as a test case. The cells were numbered sequentially from the topmost cell downward. Figure 10 illustrates the displacement of the tab area for the four layers of cells when subjected to a stack load of 2000 N.

It is evident from Figure 11 that the topmost cell exhibits the largest displacement. This is because the displacement of the upper cells is influenced by more displacement superposition factors compared to the lower cells. The primary cause of displacement within the stack is the deformation of the buffer foam cushions positioned between different cells. Additionally, Figure 10 demonstrates that the proposed method achieves displacement values that are closer to the reference values compared to the other two methods, further validating the superior accuracy of the proposed approach in this scenario.

The eight-layer cell stack was subjected to loads of 2000 N, 4000 N, 6000 N, and 8000 N. Images were captured and measurements were performed using the method described above. The calculation error and root mean square error (RMSE) of the displacement of the positive electrode lug under all load conditions were analyzed to evaluate the stability of the proposed method. Figure 12 illustrates the differences between the measured and reference values of lug displacement for all load conditions across the three methods.
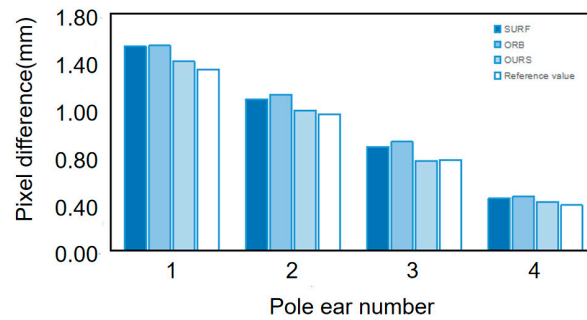
Figure 11. Displacement detection of the cell lug on each layer under a load of 2000 N.
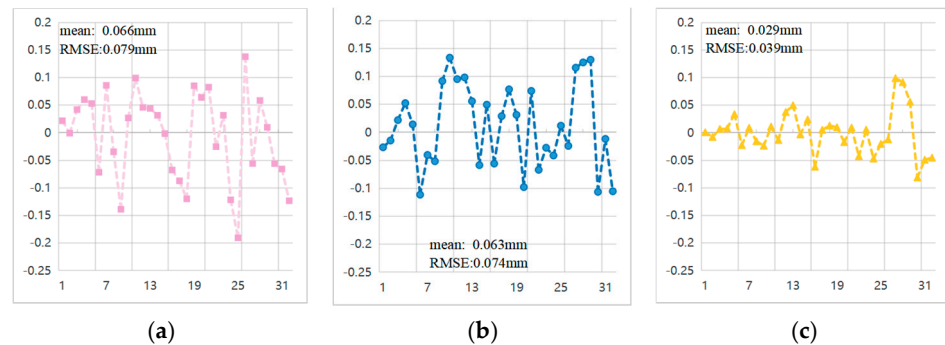

(**a**)　　　　　　　　　(**b**)　　　　　　　　　(**c**)

Figure 12. Comparison of measurement error results for lug displacement detection: (**a**) SURF−FLANN; (**b**) ORB with adaptive thresholds; (**c**) ours.

The results indicate that the average measurement error for all three methods achieves an accuracy level within 0.01 mm. However, the method proposed in this paper demonstrates superior performance, with a maximum error of only 0.09 mm and an average error of 0.029 mm. Additionally, the root mean square error (RMSE) of the proposed method is 0.039 mm, which is approximately half that of the other two methods. This highlights the higher accuracy and stability of the proposed approach.

4. Discussion and Outlook

This paper presents a displacement measurement scheme for lithium battery cells based on machine vision. The region of interest (ROI) of the cell tab is extracted using a method based on the HSV color model, enabling pixel-level displacement detection through the MicKey feature point matching process combined with a mismatching elimination approach based on the Z-Score. To determine the real-world displacement of the multi-layer cells during the assembly process, the proposed method employs a conversion factor (CF) to relate physical dimensions to pixel measurements.

To evaluate the performance of the proposed scheme, experiments were conducted to comprehensively analyze the displacement of multi-layer cells during assembly. These experiments included an accuracy test of the algorithm for a four-cell stack assembly and a stability test of multi-tab displacement detection in an eight-cell stack assembly. Based on the experimental results, the following conclusions can be drawn:

1. Image segmentation of the stack body using the HSV color model effectively produces a binary image of the cell outline with minimal noise, thereby avoiding the higher noise levels typically associated with image segmentation based on the RGB color model [27].

2. A comparison between the proposed method and laser sensor measurements indicates that the maximum absolute error is 0.08 mm, while the root mean square error (RMSE) is 0.039 mm. Compared to other existing methods, the approach proposed

in this study demonstrates superior performance in feature point detection, feature matching accuracy, and the mitigation of mismatching issues. These results confirm that the proposed method satisfies the engineering requirements for multi-layer cell displacement measurement.

In summary, the displacement measurement method proposed in this paper reliably detects and quantifies the displacement of battery cells during assembly, providing valuable support for the advancement of full health monitoring systems for lithium batteries. Moreover, this method offers a novel solution for detecting the dynamic displacement of multiple battery cells during assembly.

However, future research still faces several challenges, particularly concerning the effects of fixture vibrations and variations in ambient lighting on the imaging process. In practical assembly processes, the vibrations of the fixture can interfere with displacement measurements. Consequently, future studies should focus on enhancing the system's resilience to these interferences, potentially through the design of more stable fixtures and the optimization of imaging algorithms to minimize errors induced by vibrations. Additionally, variations in lighting conditions can significantly impact the accuracy of vision-based measurements. Future research could explore adaptive lighting adjustment technologies or conduct data training under different lighting conditions to improve the versatility of the algorithms. To ensure the reliability of measurement results, future investigations should also aim to establish standardized detection processes and methodologies to facilitate widespread adoption in various industrial environments. This includes the development of relevant testing standards and protocols to provide consistent references for the industry.

Author Contributions: Conceptualization, Y.X. (Yueda Xu) and Y.X. (Yanfeng Xing); methodology, Y.X. (Yueda Xu); software, Y.X. (Yueda Xu); validation, H.Z., Y.L. and L.R.; investigation, H.Z., L.R., Y.L. and Z.Z.; resources, Y.X. (Yueda Xu) and Z.Z.; data curation, Y.X. (Yueda Xu) and Z.Z.; writing—original draft preparation, Y.X. (Yueda Xu); writing—review and editing, Y.X. (Yueda Xu); visualization, Y.X. (Yueda Xu); supervision, Y.X. (Yueda Xu), Y.X. (Yanfeng Xing), Y.L. and L.R.; project administration, Y.X. (Yanfeng Xing); funding acquisition, Y.X. (Yanfeng Xing). All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: Hongbo Zhao is an employee of Pan Asia Technical Automotive Center Co., Ltd. Yufang Lin and Lijia Ren are employees of SAIC GM Power Technology (Shanghai) Co., Ltd. Zhihan Zhou is an employee of Avatr Technology (Chongqing) Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Lee, S.S.; Kim, T.H.; Hu, S.J.; Cai, W.W.; Abell, J.A. Joining Technologies for Automotive Lithium-Ion Battery Manufacturing: A Review. In Proceedings of the ASME 2010 International Manufacturing Science and Engineering Conference, Erie, PA, USA, 12–15 October 2010; pp. 541–549.
2. Das, A.; Li, D.; Williams, D.; Greenwood, D. Weldability and shear strength feasibility study for automotive electric vehicle battery tab interconnects. *J. Braz. Soc. Mech. Sci. Eng.* **2019**, *41*, 54. [CrossRef]
3. Zhao, Y.; Patel, Y.; Hunt, I.A.; Kareh, K.M.; Holland, A.A.; Korte, C.; Dear, J.P.; Yue, Y.; Offer, G.J. Preventing lithium ion battery failure during high temperatures by externally applied compression. *J. Energy Storage* **2017**, *13*, 296–303. [CrossRef]
4. Kang, H.; Lim, C.; Li, T.; Fu, Y.; Yan, B.; Houston, N.; De Andrade, V.; De Carlo, F.; Zhu, L. Geometric and electrochemical characteristics of $LiNi_{1/3}Mn_{1/3}Co_{1/3}O_2$ electrode with different calendering conditions. *Electrochim. Acta* **2017**, *232*, 431–438. [CrossRef]

5. Chen, W.; Han, X.; Pan, Y.; Yuan, Y.; Kong, X.; Liu, L.; Sun, Y.; Shen, W.; Xiong, R. Defects in Lithium-Ion Batteries: From Origins to Safety Risks. *Green Energy Intell. Transp.* **2024**, 100235. [CrossRef]

6. Yanhua, W.; Cheng, W.; Yan, F.; Bowen, D.; Gang, W. Application and Analyzation of the Vision-Based Structure Model Displacement Measuring Method in Cassette Structure Shaking Table Experiment. *Adv. Civ. Eng.* **2020**, *2020*, 8869935. [CrossRef]

7. 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]

8. 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]

9. Matsuya, I.; Matsumoto, F.; Ihara, I. Ultrasonic lateral displacement sensor for health monitoring in seismically isolated buildings. *Sensors* **2015**, *15*, 17000–17012. [CrossRef]

10. Du, W.; Lei, D.; Hang, Z.; Ling, Y.; Bai, P.; Zhu, F. Short-distance and long-distance bridge displacement measurement based on template matching and feature detection methods. *J. Civ. Struct. Health Monit.* **2023**, *13*, 343–360. [CrossRef]

11. Olaszek, P. Investigation of the dynamic characteristic of bridge structures using a computer vision method. *Measurement* **1999**, *25*, 227–236. [CrossRef]

12. Jáuregui, D.V.; White, K.R.; Woodward, C.B.; Leitch, K.R. Noncontact photogrammetric measurement of vertical bridge deflection. *J. Bridge Eng.* **2003**, *8*, 212–222. [CrossRef]

13. Yu, S.; Zhang, J.; He, X. An advanced vision-based deformation measurement method and application on a long-span cable-stayed bridge. *Meas. Sci. Technol.* **2020**, *31*, 065201. [CrossRef]

14. 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]

15. Feng, D. *Advanced Vision-Based Displacement Sensors for Structural Health Monitoring*; Columbia University: New York, NY, USA, 2016.

16. Feng, D.; Feng, M.Q.; Ozer, E.; Fukuda, Y. A vision-based sensor for noncontact structural displacement measurement. *Sensors* **2015**, *15*, 16557–16575. [CrossRef]

17. Zhong, X.; Yang, L.; Li, N.; Chu, Z.; Chen, J.; Zhu, S.; Song, W.-L.; Ai, S.; Chen, H.-S. In-situ characterizations and mechanism analysis of mechanical inhomogeneity in a prismatic battery module. *J. Power Sources* **2022**, *548*, 232053. [CrossRef]

18. Nam, W.; Kim, J.-Y.; Oh, K.-Y. The characterization of dynamic behavior of Li-ion battery packs for enhanced design and states identification. *Energy Convers. Manag.* **2018**, *162*, 264–275. [CrossRef]

19. Huang, L.; Li, T.; Ding, C.; Zhao, J.; Zhang, D.; Yang, G. Diagnosis of the severity of Fusarium head blight of wheat ears on the basis of image and spectral feature fusion. *Sensors* **2020**, *20*, 2887. [CrossRef]

20. Barroso-Laguna, A.; Munukutla, S.; Prisacariu, V.A.; Brachmann, E. Matching 2D Images in 3D: Metric Relative Pose from Metric Correspondences. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 17–18 June 2024; pp. 4852–4863.

21. DeTone, D.; Malisiewicz, T.; Rabinovich, A.S. Self-supervised interest point detection and description for Fisheye and Perspective Images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Salt Lake City, UT, USA, 18–23 June 2018.

22. Revaud, J.; De Souza, C.; Humenberger, M.; Weinzaepfel, P. R2d2: Reliable and repeatable detector and descriptor. *Adv. Neural Inf. Process. Syst.* **2019**, *32*. [CrossRef]

23. Tyszkiewicz, M.; Fua, P.; Trulls, E. DISK: Learning local features with policy gradient. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 14254–14265.

24. Oquab, M.; Darcet, T.; Moutakanni, T.; Vo, H.; Szafraniec, M.; Khalidov, V.; Fernandez, P.; Haziza, D.; Massa, F.; El-Nouby, A. Dinov2: Learning robust visual features without supervision. *arXiv* **2023**, arXiv:2304.07193.

25. Feng, Y.; Sun, Y. Image matching algorithm based on surf feature extraction and FLANN search. *J. Graph.* **2015**, *36*, 650–654.

26. Li, S.; Wang, Q.; Li, J. Improved ORB matching algorithm based on adaptive threshold. *J. Phys. Conf. Ser.* **2021**, *1871*, 012151. [CrossRef]

27. Zheng, L.Y.; Xu, J.X. Multi-crop-row detection based on strip analysis. In Proceedings of the 2014 International Conference on Machine Learning and Cybernetics, Lanzhou, China, 13–16 July 2014; pp. 611–614.