

3D Video Image Processing Effect Optimization Method Based on Virtual Reality Technology

Zhihong He*, Wenjie Jia, Erhua Sun, Huilong Sun

School of Digital Engineering, Chongqing College of Architecture and Technology, Chongqing
401331, China

*Corresponding author: hzhccat@163.com

Received: May 24, 2021. Revised: December 3, 2021. Accepted: January 11, 2022. Published: January 12, 2022.

Abstract—The existing optimization methods have the problem of image edge blur, which leads to a high degree of shadow residue. In order to address this problem, reduce the shadow residual degree, this paper designs a 3D video image processing effect optimization method supported by virtual reality technology. Coding was used to eliminate redundant data in video and eliminate image noise using median filtering. The virtual reality technology detects the image edge and determines the motion offset between the image frames. According to the motion parameters of the camera carrier obtained from the motion estimation, the feature point matching algorithm constructs the video image motion model, and uses the camera calibration technology to set the processing effect optimization mode. It is optimized by perspective projection transformation. Experimental results: the average shadow residual degree of the optimization method and the two existing optimization methods are 3.108%, 6.167% and 6.396% respectively, which proves that the optimization method combined with virtual reality technology has higher practical application value.

Keywords—virtual reality technology; three dimensional video image; treatment effect; optimization method.

I. INTRODUCTION

In the information society, the main way for people to obtain external information is through vision. Therefore, digital video technology has been a research hotspot in related fields. With the improvement of various encoding and decoding standards, the traditional flat video technology has become mature. In addition to getting higher definition through high-definition display devices, it does not get a better user experience of video viewing. In order to obtain a higher viewing experience, research institutions have an upsurge in the research of 3D video. Compared with traditional 2D video, 3D video provides two different video streams into the human eye, which forms the depth information of objects in the human brain and produces stereo visual experience. 3D video processing technology can be divided into two view and multi view 3D video according to the number of video streams. Dual view stereo video only provides two different video streams,

users need to wear stereo video glasses to produce stereo visual experience. Multi view 3D video, through special display technology, allows users to freely switch views and select suitable views to enjoy 3D programs. Free view stereo video processing has good application prospects in 3D video conference system, distance education, virtual reality, online shopping and other fields [1]-[3]. Free view video has the following characteristics: first, interactivity. In free view video, users can turn passive viewing into active viewing. For example, we can change our own perspective and viewpoint in a scene, and select the region of interest for viewing. Second, immersion. Users can enjoy the stereoscopic visual effect from any angle, so as to get a strong sense of the scene, and fully realize the reality of taking themselves as the protagonist in the simulation environment. Therefore, based on multi view 3D video processing technology, free view video is further developed. Video technology uses the parallax similarity between different video streams to extract the depth information of the object in the video image frame, carry out correlation coding, and only transmit the code stream of limited path video coding and the code stream of depth image coding. In the decoder, the depth image is used to reconstruct the multi-channel video stream to get the free view stereo video playback. It can be seen that video processing technology can greatly reduce the data processing bandwidth required by free view stereo video processing through depth information processing, optimization, coding and transmission, which provides a theoretical basis and implementation guarantee for the design and implementation of real-time stereo video processing system. Therefore, the depth image processing and optimization is the basis of free view stereo video processing technology. The effect of depth image processing and optimization is directly related to the quality of the final reconstructed multi view video, and ultimately affects the quality of free view stereo video viewing. Therefore, the 3D video image processing effect needs to be optimized.

Zhang et al. [4] proposed a three-dimensional dynamic visual self-optimization method based on multi-source information fusion. Three-dimensional dynamic visual image reconstruction with 3D dynamic visual image, block matching with fuzzy feature matching method, construct the image feature segmentation model of three-dimensional dynamic visual image, realize the definition of 3D dynamic

visual image, and self-optimization processing in multi-source information fusion space.

Zoican [5] proposed an optimized denoising method for video image editing effect based on Blackfin. Identify the noise pixels in the video clip frame; filter the median image pixels for each frame; determine the remaining time before the operation of each processing stage, and reduce the number of runs in each stage; and optimize the 3D dynamic image combined with the QoS scheduling algorithm and Blackfin.

Due to the poor video coding method, the processed image edges are blurred to a high degree of residual shadow. To reduce the shadow residual, propose optimization of 3D video image processing based on virtual reality. Video and image coding methods are used to eliminate the redundant data in the video, and the median filtering method is used to eliminate the image noise; Based on virtual reality technology, the image generated by the static imaging system is represented by a matrix, detecting the fuzzy edges for extracting image motion feature points with Harris operator, match feature points with the mean clustering algorithm, build the video image motion model, eliminate the offset vector from the unintentional motion of the video image, and Intel parallel acceleration library is used to accelerate the calculation of the algorithm; The camera calibration technology is used to splicing images according to the real image sequence, organized into panoramic views, and optimized by perspective projection transformation. The experimental results prove that the paper reduces the shadow residual degree, clarifies the fuzzy image edge, improves the 3D video image quality, and thus has high practical value.

II. 3D VIDEO IMAGE PROCESSING EFFECT OPTIMIZATION METHOD BASED ON VIRTUAL REALITY TECHNOLOGY

A. Median filtering for 3D video image acquisition

The 3D video images contain large amounts of noise and redundant data that need to be preprocessed. Video and image coding are used to eliminate redundant data in video, and the median filtering method is used to eliminate image noise.

Since digital video began to replace analog video, digital video applications have become the most important driving force of the information industry. Compared with simulated video, digital video has the advantages of convenient editing, good reproducibility and long transmission distance. But the huge amount of data presents a severe test for the limited bandwidth of the transmission system. Therefore, the compression of video data makes the possible transmission and storage of large amounts of video data. Among the many branches of digital video research, video coding is one of the most challenging branches. The aim of video coding research is to eliminate substantial redundancy in digital video, thereby saving transmission and storage costs. During digitization, oversampling of simulated video signals is the fundamental cause for the large amount of data redundancy in digital video [6], [7]. The key to video coding is to eliminate various redundancy

in video information. A large number of statistics suggest a correlation between adjacent pixels within the same frame and with little likelihood of mutation. Since the residual value is much less than the pixel value itself, the purpose of encoding compression is achieved. Depending on the type of redundancy, prediction codes can be divided into internal and internal predictions [8]-[10]. Interframe prediction is also known as motion prediction. Each pixel yields the corresponding predicted values based on the motion vector and reference, used to eliminate the spatial redundancy of the image. The main purpose of image coding and compression is to eliminate the spatial correlation of the image data. The goal of eliminating spatial redundancy is to reduce the interpixel correlation by transform encoding. The image transformation itself does not compress the data, but makes the image data in the transformation domain the least coherent. The DC and low-frequency regions in the image represent most of the image. By transforming coding, images of the spatial domain can be transformed to other frequency domains [11]-[13]. Using the characteristics of human visual insensitive to high frequency signals, a rough quantification of the high frequency coefficient can be used to better eliminate psychological redundancy.

The general method blur the image when filtering noise, thus losing some edge information of the image. In this paper, the median filtering method is used for image preprocessing, which is to eliminate the noise in the image without damaging the image details, if not handled, will affect the subsequent operation of the image. The median filtering for the image is as follows:

$$p(a,b) = \frac{1}{w \cdot e} \sum_{e=1} (c) \quad (1)$$

In Formula (1), w is the average value of image pixels, e is the set of pixel positions, and c is the image noise.

B. Image edge detection based on virtual reality technology

Virtual reality is an advanced user interface. In the past, it usually refers to a kind of three-dimensional reality realized by a series of sensing auxiliary facilities such as helmet mounted display and sensing gloves. Through these facilities, people send various motion information to the computer in a natural way (such as the rotation of the head, the movement of the body, etc.), and get the computer's visual, auditory, tactile and other sensory feedback to the user. These sensory feedback changes with people's different movements. The purpose of virtual reality technology is to create a harmonious man-machine environment which is similar to the objective environment but beyond the objective time and space, and can be immersed in and controlled by it, that is, a controllable space composed of multi-dimensional information. Its most important goal is real experience and convenient and natural human-computer interaction. Immersion refers to the degree of reality that users feel as protagonists in the virtual environment. Ideally, the virtual environment should make it difficult for users to distinguish the true from the false, and the visual scene should change with the user's point of view, so that users have the feeling of being on the

scene [14]. For example, users can directly grasp objects in the virtual environment with their hands. Digital video sequence is composed of one frame of image. The high-frequency noise component will be introduced into the video image, so that there is a difference between the frames of the image. We regard the change between the frames as motion. The former deals with the whole image sequence, while the latter deals with the blurred image frame by frame. Video blur can be divided into intra blur and inter blur, which are caused by intra motion and inter motion. Since digital image is obtained by numerical sampling of analog image, we can use matrix to represent the image generated by static imaging system, and the elements in the matrix are the values of sampling points. Then the expression formula of motion estimation vector of video frame is as follows:

$$w = g(i) + D(i) \quad (2)$$

In Formula (2), g represents the low frequency component in the camera, D represents the high frequency component in the camera, and i represents the compensation vector of the i -th frame. Imagination means that users immerse themselves in multi-dimensional information space, rely on their own perception and cognitive ability to acquire knowledge in an all-round way, play their subjective initiative, seek answers and form new ideas. To achieve the stability of image sequence is the development trend of image stabilization technology. Some algorithm is selected to process the image, determine the motion offset between the image frames, separate the random jitter in the global motion of the image, retain the camera's intentional motion in the horizontal and vertical directions, and compensate the offset in the motion direction, so as to achieve the purpose of image stabilization. First of all, preprocessing is carried out for each input image in the video sequence, which is mainly used for image enhancement and denoising to eliminate unnecessary interference noise and improve image stabilization accuracy. Secondly, the problem of motion deviation between each frame is dealt with. The global motion parameters of the image sequence are obtained by motion estimation and evaluated comprehensively to distinguish intentional motion (such as normal scanning motion of the camera) and unintentional motion (such as rotation, zoom and random jitter of the camera). The unintentional components are separated and brought into the motion model, the global motion components that need to be compensated between each frame of the image are calculated accurately, including the rotation angle and zoom ratio of the image, the horizontal and vertical displacement, etc. Finally, the motion compensation method is used to remove the unintentional component of the motion between the frames of the image sequence, and retain the intentional component in the image stabilization system, so as to obtain a smooth and continuous output image sequence. Motion estimation is one of the key technologies in the electronic image stabilization algorithm. By determining the offset of the local motion of the image, the global motion vector of the image sequence can be judged. Its speed and accuracy will affect the stability of

the system [15]. Motion decision is another difficulty in the electronic image stabilization algorithm that is to separate the unintentional motion and intentional motion from the motion components, get the translation, rotation and scaling vectors among the images, and compensate the motion compensation module of the image, reorganize the image rows and columns, and output a stable image.

C. Video image motion model based on feature point matching algorithm

Using the motion compensation method to remove unintentional components of inter-frame motion in the image sequence and retain the intentional component requires estimation of the global motion component. To this end, this paper constructs a video image motion model based on the feature point matching algorithm to estimate the motion components. Image motor feature points were extracted using the Harris operator; The mean clustering algorithm matches the video image motion model and stabilizes the image; According to the offset vector obtained by the motion estimation module, the unintentional motion of the video image was eliminated and the position of the image was corrected; Intel parallel acceleration library is used to speed up the above algorithm.

The base feature point is the corner point in the image containing rich information in detail, which can reflect the shape and local characteristics of the image. In images, we can easily locate angles, especially in artificial target scenes, and in computer vision, angle extraction algorithms are also relatively mature [16]. The Harris operator has its own advantages in stability and robustness over other operators [17], [18]. The high function of the Harris operator can reduce the noise component compared to the binary component. Because the Harris operator is only based on the first-order differential and Gaussian filtering of the image gray scale, so the real-time and stability are strong, but it is not sensitive to the image gray scale change, rotation, etc. When the gray scale of the image is too low (the image is very dark), only the number of corner points cannot be extracted or too small to correct the image, so it is necessary to increase the image contrast to increase the number of angles.

Angle point matching is required after angle extraction. After angular matching, equivalent mismatching pairs occur, which need to be eliminated by the mean clustering algorithm. If the multiplicative transformation matrix is too large, indicating that the transformation has been found in the current scene, the transformation matrix needs to be reset until the angular match is completed.

Since the algorithm integrates multiple modules, the general computational program cannot meet the real-time requirements, so the image algorithm adopts Intel's parallel acceleration library, and adopts graphics card acceleration to speed up the computation.

The calculation formula of Harris corner detection operator is as follows:

$$\frac{\alpha h}{\alpha u} = h(u+1, q) - h(u, q) \quad (3)$$

In Formula (3), h is the image window, u is the

image gray level, α is the image offset, and q is the image corner response function. When a window moves in any direction in the flat area of the image, the gray image window in each direction will not change. When the window moves along the edge, the gray level of the image within the window does not change, but when the window moves in any direction of the image corner, the gray level of the window changes obviously, so we can determine whether it is a corner according to the degree of change of the window in each direction [19]. Due to the high matching accuracy, small amount of calculation and strong adaptability of image feature points, they have been widely used in many fields. Compared with block matching algorithm, the feature points extraction of motion vector estimation can reduce the amount of calculation, improve the speed of calculation and provide a reliable image matching. The compensation of the image is the final step of stabilizing the image. According to the offset vector obtained from the motion estimation module. On the basis of Formula (3), the expression formula of video image motion model is obtained as follows:

$$\begin{bmatrix} u_i \\ q_i \end{bmatrix} = \frac{1}{y} \begin{bmatrix} 1 & -\alpha \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} u' & ku \\ q' & kq \end{bmatrix} \quad (4)$$

In Formula (4), u is the image gray level, α is the image offset, q is the image corner response function, y is the compensation motion vector, and k is the mapping coordinate value. The video image acquired by the camera includes unintentional motion (translation, rotation and zoom of the camera carrier) and intentional motion (normal scanning of the camera). Unintentional motion of video images was removed according to motion estimation and conscious motion was retained to correct the position of the image.

D. Camera calibration technology setting processing effect optimization mode

As a kind of vision sensor, camera is the necessary equipment of virtual scene construction and virtual reality interaction technology, and the direct source of visual information. Camera geometry model parameters will have an important impact on panorama generation and hand tracking. Camera parameters are generally divided into external parameters. Camera calibration has important application value in computer vision, scene reconstruction, motion estimation and virtual reality. In this project, the construction of light flight simulator extravehicular scene based on image rendering is based on a group of real image sequences captured by the camera. Image processing technology is used to splice them and organize them into panorama. In the whole process, the focal length of the camera needs to be determined first [20]-[22]. If the calibration is not accurate, the panorama will be distorted, misplaced and other phenomena. Virtual reality finger location technology based on computer vision uses camera as tracking equipment, camera calibration is an important step of finger location. Among them, the internal parameters of the camera have an important impact on the positioning accuracy, and the determination of the external

parameters of the camera provides a technical solution for finger positioning. The optimized region of image and video is determined by camera calibration technology, as shown in Figure 1:

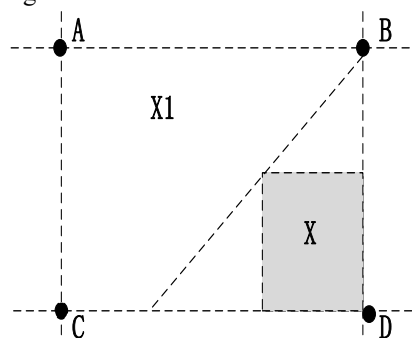


Figure 1. Camera calibration area

It can be seen from Figure 1 that X is the camera calibration area, and a variety of different camera calibration methods are proposed from different angles. From the perspective of realization principle, there are linear calibration and nonlinear calibration methods; It can be divided into non self-calibration and self-calibration according to whether the reference object needs to be calibrated; From the point of view of calibration steps, it can be divided into two steps, three steps, four steps and so on; From the perspective of technology progress, there are traditional calibration technology, self-calibration technology and calibration technology based on active vision. The traditional calibration method needs to use the precision processed calibration block. The vertex of each small square on the calibration block is used as the feature point. The process can be described by imaging transformation. Because most of the images are obtained by perspective projection camera, perspective projection transformation is used to optimize. Then the expression formula of the proportional relationship between the optical center and the projection point is as follows:

$$\begin{cases} J' = \frac{t \cdot s + 1}{Z} \\ R' = \frac{t \cdot s - 1}{Z} \end{cases} \quad (5)$$

In Formula (5), t is the focal length of the camera, s is the projection position of any point in space, and Z is the scale factor. The process of imaging transformation mainly involves coordinate transformation between different coordinate systems. The general 3D scene is represented by this coordinate system. According to the different transformation relations between the above coordinate systems, different imaging geometric models can be obtained, also known as camera imaging models. Camera imaging models mainly include linear model and nonlinear model. Among them, the nonlinear model considers that the lens is not ideal perspective imaging and has different degrees of distortion, so the nonlinear model camera calibration considers distortion parameters and introduces nonlinear optimization, which is cumbersome, slow, sensitive to initial value selection and noise, and the nonlinear search can not guarantee that the parameters

converge to the global optimal solution; Pinhole imaging model, also known as linear camera model, has the characteristics of simple and accurate. As a general camera model, it is widely used in computer vision and other fields.

III. EXPERIMENTAL ANALYSIS

This paper designs an optimization method of 3D video image processing effect based on virtual reality technology. In order to test the actual effect of the optimization method, the experiment is carried out.

A. Setting up experimental environment

In the experiment, two existing optimization methods are selected and compared with the optimization method of this design. According to the needs of experimental test, the experimental environment is built as shown in Table 1:

Table 1. Experimental configuration and parameters

Configure	Parameter	Unit
Full frame rate	31	Hz
Pixel size	4.63×4.63	um
Pixel clock	66	MHz
Line rate	33.154	KHz
Number of data bits	8	bit
Sensitivity	1.32	lux
Gain	14	db
Time of exposure	0.054	ms
Voltage	16.1	V
Electric current	0.24	A

On the basis of Table 1, the experimental test is carried out and the experimental results are obtained.

B. Experimental result

Under different resolution image conditions, the shadow residual degree of the optimization method and the two optimization methods are tested respectively. The lower the value is, the better the performance of the optimization method is. The experimental results are shown in Tables 2-4:

It can be seen from Table 2 that the average shadow residual degree of the optimization method and the two existing optimization methods are 3.257%, 3.846% and 1.078% respectively; It can be seen from Table 3 that the average shadow residual degree of the optimization method and the two existing optimization methods are 6.074%, 6.882% and 3.169% respectively; It can be seen from Table 4 that the average shadow residual degree of the optimization method in this design and the two existing optimization methods are 9.170%, 8.457% and 5.076% respectively, indicating that the optimization method in this design has better performance. This is because this paper uses virtual reality technology to detect the image edges, laying the foundation for fuzzy edge clarity; using the Harris operator to obtain the camera motion offset vector, eliminate the unintentional motion of the video image, correct the image position; finally using the perspective projection conversion, so the optimization effect is high.

Table 2. Resolution 1920 × 1080 shadow residual degree (%)

Number of experiments (Group)	Optimization method 1	Optimization method 2	Optimization method of design
1	3.224	4.112	1.102
2	2.641	3.674	0.997
3	3.105	3.584	1.113
4	3.668	3.621	0.899
5	2.884	4.112	1.224
6	4.021	3.991	1.136

Table 3. Resolution 1280 × 720 shadow residual degree (%)

Number of experiments (Group)	Optimization method 1	Optimization method 2	Optimization method of design
1	6.612	6.471	3.121
2	5.879	7.101	3.504
3	6.103	6.645	2.997
4	5.968	7.206	3.232
5	6.036	6.668	3.058
6	5.846	7.204	3.107

Table 4. Resolution 800 × 600 shadow residual degree (%)

Number of experiments (Group)	Optimization method 1	Optimization method 2	Optimization method of design
1	9.945	7.748	5.102
2	8.978	8.554	5.211
3	9.756	9.023	4.996
4	7.889	7.665	5.061
5	8.766	8.447	4.987
6	9.687	9.306	5.102

IV. CONCLUSION

In order to reduce the shadow residual of 3D video image edges, the optimization method of 3D video image processing effect based on virtual reality technology is proposed. Eliminate redundant data by video and image coding, eliminate image noise; detect image edge based on virtual reality technology, extract image motion feature points, construct video image motion model; obtain the offset vector, correct the image position; use parallel acceleration library to accelerate the calculation speed of the algorithm; splicing image and optimize the image fuzzy edge. Through experimental testing, we show that the proposed method has better use than two existing optimization methods, making the 3D dynamic images clearer. At the same time, the academic literature of virtual reality technology and optimization methods is enriched. Due to the limitations of the research conditions, this paper has conducted no in-depth research on the application of VR technology in other fields, which needs to be further improved.

REFERENCES

- [1] F. Zhang and F. Wang, "Exercise fatigue detection algorithm based on video image information extraction," IEEE Access, vol. 8, no. 2, pp. 199696-199709, 2020.

- [2] J. Dai, P. P. Zhang, D. Wang, H. C. Lu and H. Y. Wang, "Video person re-identification by temporal residual learning", *IEEE Transactions on Image Processing*, vol. 28, no. 3, pp. 1366-1377, 2019..
- [3] H. Wang, Y. Zhang and X. Fan, "Rapid early fire smoke detection system using slope fitting in video image histogram", *Fire Technology*, vol. 56, no. 2, pp. 695-714, 2020.
- [4] G. H. Zhang, B. Feng, Q. Bo and L. Yu, "3D dynamic scene self-correction simulation based on multi-source information fusion", *Computer Simulation*, vol. 36, no. 11, pp. 427-431, 2019.
- [5] S. Zoican, "QoS scheduling algorithm for videoclips denoising", *WSEAS Transactions on Signal Processing*, vol. 17, pp. 41-45, 2021..
- [6] P. Shamsolmoali, M. E. Celebi and R. Wang "Advances in deep learning for real-time image and video reconstruction and processing", *Journal of Real-Time Image Processing*, vol. 17, no. 6, pp. 1883-1884, 2020.
- [7] A. Brifman, Y. Romano and M. Elad, "Unified single-image and video super-resolution via denoising algorithms", *IEEE Transactions on Image Processing*, vol. 28, no. 12, pp. 6063-6076, 2019.
- [8] A. Ku. Bhunia, A. Konwer, A. K. Bhunia, A. Bhowmick, P. P. Roy and U. Pal, "Script identification in natural scene image and video frames using an attention based Convolutional-LSTM network", *Pattern Recognition*, vol. 85, no. 8, pp. 172-184, 2019.
- [9] D. Kim, D. Joo and J. Kim, "TiVGAN: Text to image to video generation with step-by-step evolutionary generator", *IEEE Access*, vol. 8, pp. 153113-153122, 2020.
- [10] L. Krasula, B. Yoann and P. L. Callet. "Training objective image and video quality estimators using multiple databases", *IEEE Transactions on Multimedia*, vol. 22, no. 4, pp. 961-969, 2019.
- [11] H. Prasetyo, C. H. Hsia and C. H. Liu "Vulnerability attacks of SVD-based video watermarking scheme in an IOT environment", *IEEE Access*, vol. 8, no. 1, pp. 69919-69936, 2020..
- [12] A. M. Atto, A. Benoit and P. Lambert, "Timed-image based deep learning for action recognition in video sequences", *Pattern Recognition*, vol. 104, no. 9, pp. 107353, 2020.
- [13] A. Lucas, A. K. Katsaggelos, S. Lopez-Tapuia and R. Molina, "Generative adversarial networks and perceptual losses for video super-resolution", *IEEE Transactions on Image Processing*, vol. 4, no. 9, pp. 3312-3327, 2019.
- [14] I. Kajo, N. Kamel and Y. Ruichek, "Self-motion-assisted tensor completion method for background initialization in complex video sequences", *IEEE Transactions on Image Processing*, vol. 29, pp. 1915-1928, 2019.
- [15] S. Kim and D. Y. Suh, "Recursive conditional generative adversarial networks for video transformation", *IEEE Access*, vol. 89, no. 1, pp. 37807-37821, 2019.
- [16] L. Hakim and M. I. Zul, "Implementation of discrete wavelet transform on movement images and recognition by artificial neural network algorithm", *WSEAS Transactions on Signal Processing*, vol. 15, pp. 149-154, 2019.
- [17] H. Y. Suen, K. E. Hung and Y. S. Su, "Predicting behavioral competencies automatically from facial expressions in real-time video-recorded interviews", *Journal of Real-Time Image Processing*, vol. 1, no. 27, pp. 1587, 2021.
- [18] S. I. Cho and S. J. Kang, "Temporal incoherence-free video retargeting using foreground aware extrapolation", *IEEE Transactions on Image Processing*, vol. 29, pp. 4848-4861, 2020.
- [19] G. K. Gultekin and A. Saranlı, "Multi-frame motion deblurring of video using the natural oscillatory motion of dexterous legged robots", *IET Image Processing*, vol. 13, no. 9, pp. 1502-1508, 2019.
- [20] M. Haghighat, R. Mathew, A. Naman and D. Taubman, "Illumination estimation and compensation of low frame rate video sequences for wavelet-based video compression", *IEEE Transactions on Image Processing*, vol. 28, no. 9, pp. 4313-4327, 2019.
- [21] P. S. Sivakumar, K. Kagawa, C. Crouzet, B. Choi and S. Kawahito, "Multi-exposure laser speckle contrast imaging using a video-rate multi-tap charge modulation image sensor", *Optics Express*, vol. 27, no. 18, pp. 26175-26191, 2019, .
- [22] F. Yuan and J. He, "Human height measurement in surveillance video based on vision technology", *International Core Journal of Engineering*, vol. 6, no. 5, pp. 198-208, 2020.

**Creative Commons Attribution License 4.0
(Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US