Extracting and Reconstructing Human Behavior Video Feature in High Voltage Exclusion Zone

Jun Chen\textsuperscript{1,2}, Jian Sun\textsuperscript{2}, Yuanliang Rao\textsuperscript{1} and Xiaoyong Rao\textsuperscript{2}

\textsuperscript{1}State Grid Fuzhou, Dongxiang Power Supply Branch, Jiangxi, 344000.
\textsuperscript{2}College of Electrical Engineering & New Energy, China Three Gorges University, Yichang, 443002.
Email: sunjian2009@ctgu.edu.cn

Abstract. A method is presented to the discrimination, extraction and reconstruction of human behavior in the area forbidden by high voltage, in particular, a method of scene reproduction of the accident scene when a high-voltage area endangers the power equipment injury or causes a personal property accident due to human behavior. To validate this method, an experiment was conducted on a small lake with high voltage transmission line over through. The Server video after the proposed method can be used to recognize human behavior, and transmission efficiency has been improved up to 35%.

1. Introduction
In the power transmission and distribution system, the voltage level is usually medium-high, and most overhead lines are bare. Therefore, it is required that the charged body should have a safe net distance in space, and there is a certain range of prohibited areas near the high-voltage lines to prevent personal injury and ensure the normal operation of power grid[1-2].

In recent years, the personal casualties caused by activities in the high-voltage prohibited areas have not only brought suffering and helplessness to the families of the injured, but also brought hidden dangers to judicial disputes and public images of power supply enterprises[3-4]. How to effectively prevent occurrence of personal casualties and real reappearance by scene video is a problem for safety management of power supply enterprises.

In addition, human casualties caused by human behavior in high-voltage forbidden areas mainly occur in power distribution systems. In the existing online power monitoring system, a number of video compression techniques, such as compression perception theory, have been adopted[6,7,8]. However, the amount of data transmitted from casualty site to management department is still too large, which is harmful to efficiency of communication system in distribution network. Therefore, extracting the important characteristics of human behavior for transmission is one of the core elements of reducing the distribution network communication burden.

In addition, traditional video monitoring systems are generally man-made. Due to limits of human physiology, attention will be reduced by 70% after three hours of continuous viewing[9-10]. Filtering out video information users does not care about and extracting key information from a large amount of data can reduce labor intensity, and can also improve alarm timeliness.

A method of human behavior recognition, feature image extraction and scene video reconstruction of prohibited area with high voltage is proposed in this paper, which firstly uses a detection system to scan and capture moving objects entering the forbidden area. If the capture is successful, the video device is turned on to record the object entering the forbidden area, the generated image information is transmitted to the controller, and the controller completes the rapid recognition of human behavior and
the extraction of feature images, then select the target object motion trajectory and other related information. The extracted information is transmitted to backend server, and finally the scene is reconstructed and reproduced according to the information sent by the controller.

This paper is organized as following: in Section 2, the complete proposed method for human behavior feature extraction and reconstruction from video in prohibited area with high voltage is described; then Section 3 depicts the test data and experiment of video reconstruction with the analysis of result; finally the discussion and conclusion is in Section 4.

2. Methodology
The basic principle structure diagram is shown in Figure 1.

![Diagram](image)

**Figure 1.** Schematic diagram of rapid recognition, extraction and reconstruction of human behavior in prohibited area with high voltage.

The regional parameter initialization module is divided into regional modules according to the surface characteristics of the high-pressure forbidden area as a full-background image of the high-pressure forbidden area, as shown in Figure 2. In the area diagram shown in figure II, the size of the area depends mainly on whether there are significant differences in surface characteristics, that is, the choice of significant differences in the surface as the boundary of the adjacent area. At the same time, GPS is used to complete clock proofreading and the CCD coordinate system is established as the origin of the video recorder to generate regional parameters. After the background image division and parameter setting are completed, it is stored in the controller to form a backend image database.

2.1. Human Behavior Feature Image Extraction
The goal of human behavior is usually large. Therefore, the clustering algorithm based on the contour coefficient is used to adjust the threshold of contour characteristics to realize whether it is human behavior. At the same time, the number of invading objects is determined by the number of outlines. When calculating the trajectory of each invading object, the target object is reduced to a particle. According to the change of the particle in the CCD coordinate system(X, Y) parameter, the identification of the behavior pattern of the tracking object, that is, the invading target, is realized. In tracking the target's safe distance detection, the image is first normalized with gamma/color[11]. The image is refined into black. Each black is then differentiated into several cells. The gradient amplitude and direction at each point are calculated, and the cell features are connected in series. Form a HOG feature vector. Calculate the safe distance between the line and the target object.

The recognition of video image features is mainly based on the idea that 3D video can be represented by a series of collections of ordered stationary attitude pictures. These ordered static attitude feature images can be identified and extracted by the characteristics of frame space domain and time domain[12]. That is, the static attitude of a single feature image is obtained by the
relationship between the nodes and the spatial domain features. The motion vector model is constructed by the relation between the position of the adjacent feature images before and after the same joint, and the attitude time domain features are formed. The weight matrix corresponding to each key node and attitude is calculated using a bilinear classifier. It can be set that human behavior includes K human nodes, N external nodes connected to human nodes, and M human behavior (large mechanical objects) nodes. Namely:

\[ M = \{ p_1, p_2, L, p_k, w_1, w_2, L, w_k, j_1, j_2, L, j_k \} \] (1)

The spatial coordinates of the first node are expressed as: then the spatial domain features can be expressed as:

\[ f_M = \{ m_i - m_j \mid i, j = 1, 2, L, k, n, m; i \neq j \} \] (2)

Set the moment when the video image is defined as a feature image and extracted. In addition, the video image feature image should also contain the first picture at the time of shooting and the video image at the time of still time interval in still mode; In addition, the intersection of each sub-region of the movement trajectory of the invading object is also defined as the extraction of characteristic images.

2.2. Reconstruction Based on Human Behavior Images

The feature pictures and background gray parameters, particle motion and trajectory equations are transmitted to the receiving device with decoding via a single communication mode via a coded transmitter. The receiver completes the restoration of feature pictures, background pictures; target trajectory functions, etc., and passes the restored information to the video reconstructor.

The video refactor is placed in the background server, and the reconstructor has a built-in refactoring algorithm. Its main function is to receive the receiving device to the stationary human posture image (external change) and read the motion vector of the same node from the picture of two adjacent moments. That is, in two adjacent feature graphs with time-domain attributes, the attitude timing dynamic characteristics of the same node motion vector can be expressed as [13]:

\[ f_M(i) = \{ m_i^{t+1} - m_i^t \mid m_i^t \in m_i; m_i^{t+1} \in m_{i+1}; i = 1, 2, L, k, n, m \} \] (3)

The target trajectory equation is: the orthogonal Gaussian matrix formed by the free eigenvalue measurement. The transitional image reconstruction of the video background is based on two feature images as reference images, and centered on the current image to be reconstructed. The search image block is used as the entire assumption of the current connection image. Obtained by using the Tikhonov[14] regularization method, the expression is:

\[ X_{ref(i)} = \sum_{m=1}^{N} \left\{ \begin{array}{ll}
\frac{\exp(-\frac{e_i^m}{c})}{\sum_{m=1}^{k} \exp(-\frac{e_i^m}{c})} \times x_{l(i)}^m, & e_i^m \leq T_u \\
\min\left(\frac{\exp(-\frac{e_i^m}{c})}{\sum_{m=1}^{k} \exp(-\frac{e_i^m}{c})}, \frac{\exp(-\frac{e_i^m}{c})}{\sum_{m=1}^{k} \exp(-\frac{e_i^m}{c})}\right), & e_i^m > T_u \end{array} \right. \] (4)

In the formula (4): the difference between the color and difference of the image subblock, the difference threshold, and the image subblock with time-domain properties.
Through the above process, a series of dynamic relationships with time changes are constructed, thus transitioning to a series of images with time domain characteristics. The series of images generated by the reconstructed algorithm are played on the generated time series images at 60-100 Hz frequency, and the live scene video is restored by the terminal device.

3. Data material and Experiment
The method of discriminating and extracting human behavior in the high-voltage forbidden area and scene video reconstruction is mainly composed of two parts, the front and background, and the flow chart of the realization principle is shown in Figure. 2 and Figure. 3 respectively. In the front section, the video recorder takes a panoramic view of the high-pressure forbidden area at the beginning, and the area is divided by the network density clustering algorithm to form several sub-areas. The CCD coordinate system determines the relevant parameters of each sub-area, and the GPS module completes the clock proofreading.

In feature image recognition and extraction, video indexes are created for video recordings, and images are extracted according to a certain time window, and compared with the key feature parameters of the previous image. Such as background area differences, changes in the attitude of the target object, and changes in the shape(or size) of the handheld object. If it happens, the image is judged to be a key feature image and extracted. Otherwise, the image is considered to be a similar picture of the previous picture, that is, the picture can be obtained by a certain rule based on the previous picture. At the same time, the space parameters generated by the object of the video history picture fit the trajectory of the object, and the trajectory is classified as a feature picture category as important information.

After the feature picture (information) is encoded according to certain rules, the transmitter transmits the relevant information through the information transmission channel; the receiver receives the relevant feature information and passes it to the background manager. The background manager receives the information based on the confirmed correct information, restores it to the original sent feature image with feature parameters, and according to the target's trajectory and the gray scale(or attitude change) of the same area feature image(or front and rear adjacent feature image), According to the trend algorithm, the image reconstruction between two adjacent feature images is completed, and the event scene is reconstructed at the end of the time.

![Figure 2. Map of background image division of high voltage prohibited area.](image_url)
Figure 3. Information extraction and event scene video reconstruction flow chart in background.

Figure 4. Comparison of original frame and reconstructed frame.

4. Conclusions
Extracting and reconstructing human behavior video feature is still a challenging problem for the computer vision, photogrammetry and remote sensing, etc. In this paper, a method of human behavior recognition, feature image extraction and scene video reconstruction of the high-voltage exclusion area is proposed, we firstly build a detection system to scan and capture moving objects entering the forbidden area. If the capture is successful, the video device is turned on to record the object, and the generated image information is transmitted to the controller, and the controller completes the rapid recognition of human behavior and the extraction of feature images. Selection of the target object motion trajectory and other related information. The extraction information passes through the transmission channel to the background servers, where images can be reconstructed according to the information sent by the controller, and the scene is reproduced.

The experimental results show that the proposed method has better performance in image reconstruction quality and reconstruction time, and are robust to noise.
Additionally, the equipment and method constructed in this paper can be explored to other outdoors measurements, such as high-voltage line snow covering detection, and so on.

5. Acknowledgments
This work was supported by Foundation Project (5218F0180047) of State Grid Fuzhou, Dongxiang Power Supply Branch. We also appreciate Prof. Wang Qiang, who gave some instructions and comments.

6. References
[1] ZHOU Lin-lin,HU Xiao-jun,ZHANG Lu-yin,et al. Method of Person Identity Attributes Based on Faster Region Convolution Neural Network. Journal of Electronic Components and Information Technology, 2017, 1(4):1-5.
[2] LI De-wei. Development and Application of Intelligent Operation and Maintenance Monitoring Platform Based on Multidimensional Data. Journal of Electronic Components and Information Technology, 2018, No.14 (08): 41-44.
[3] Wu,X.,Xu,D.,Duan,L.,Luo,J. Action Recognition Using Context and Appearance Distribution Features.In:CVPR(2011).
[4] XIA Li-min,YANG Bao-juan,TU Hong-bin. Recognition of Suspicious Behavior Using Case-Based Reasoning. Journal of Central South University, 2015, 22(1):241-250.
[5] H.Wang,A.Klaser,C.Schmid. Dense trajectories and Motion Boundary Descriptors for Action Recognition. IJCV, 2013, 103(1):60-79.
[6] Sandhan,T.,Choi,J.Y. Frequency Grams and Multi-Feature Joint Sparse Representation for Action and Gesture Recognition.In:ICIP(2014).
[7] Laptev I,Marszalek M,Schmid C,et al. Learning Realistic Human Actions from Movies[C]. Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015:1-8.
[8] Laptev I,Marszalek M,Schmid C,et al. Learning Realistic Human Actions from Movies[C]. Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015:1-8.
[9] Wang H,Klaser A,Schmid C,et al. Dense Trajectories and Motion Boundary Descriptors for Action Recognition. International Journal of Computer Vision, 2013, 103(1):60-79.
[10] Jargalsaikhan I,Little S,Trichet R,et al. Action Recognition in Video Using a Spatial-Temporal Graph-Based Feature Representation . AVSS 2015.
[11] Rudy Cahyadi H.P,Junaidillah Fadli,Hsing-Kuo Pao. Group Behavior Recognition Based on Dictionary and Hierarchical Learning. Proceedings of Computer Science, 2018,144.
[12] Diaz-Pereira,M. Pino,Gomez-Conde,Ivan,Escalona M,et al. Automatic Recognition and Scoring of Olympic Rhythmic Gymnastics Movements. Human Movement Science, 2014, 34:63-80.
[13] Ahad M A R,Tan J K,Kim H,et al. Motion History Image: its Variants and Applications.Machine Vision and Applications,2012,23(2):255-281.
[14] Sadanand S,Corso J J,Action Bank. A High-Level Representation of Activity in Video,Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012.
[15] Cai,Q.,Yin,Y.,Man,H. Learning Spatio-Temporal Dependencies for Action Recognition. In:ICIP(2013).
[16] Wu, X., Xu, D., Duan, L.,Luo,J. Action Recognition Using Context and Appearance Distribution Features. In:CVPR(2011).