Amended Claims

1. As an ~~UAV~~ unmanned aerial vehicle ("**UAV**") inspection method for power line based on human visual system, it has such characteristics and includes the following steps:

(1) UAV video capture: UAV armed with binocular visual measuring equipment is used to obtain the video image of distribution of the power line and the environmental information about the power line;

(2) The image preprocessing module grabs the frame of the video image sequence of the power line input system and preprocess the current frame of image, including graying processing of image and ~~DoG~~ Difference of Gaussian ("**DoG**") edge detection; the DoG edge detection as described uses the DoG results of different parameters to describe the image edge and applies the same treatment to the left visual image and right visual image in sequence;

(3) The power line detection module uses the mathematical morphology to deal with the image that has been preprocessed, selects the structural factors in the same direction as the power line to provide the image with repeated expansion and corrosion treatment to remove the noise in the image, and then uses the human visual connected domain attention mechanism to select the largest linear connected domain to complete segmentation of the power line in image;

Separately record the horizontal coordinate positions ($x_{dz1}, x_{dz2}...x_{dzj}$) and ($x_{dy1}, x_{dy2}...x_{dyj}$) of the power line in left visual image and right visual image; where $x_{dz1}, x_{dz2}...x_{dzj}$ and $x_{dy1}, x_{dy2}...x_{dyj}$ separately refer to the horizontal coordinate of the central point in the left visual image and right visual image of the connected domain of *j* power lines;

(4) The binocular image registration module provides registration for the left

visual image and right visual image separately and uses ~~SURF~~ Speed Up Robust Features ("**SURF**") algorithm to find the feature points of the left and right visual edge images which have been preprocessed and obtain the descriptor of the current feature point, then provides exact match of the feature points and record the location information of the exact matching point in left and right visual images （x_{z1}, y_{z1}）~（x_{y1}, y_{y1}）,（x_{z2}, y_{z2}）~（x_{y2}, y_{y2}）...（x_{zn}, y_{zn}）~（x_{yn}, y_{yn}）, of which *n* refers to the number of all feature points included in single image, x_{zn}, y_{zn} and x_{yn}, y_{yn} refer to the position coordinates of corresponding features points in the left and right visual images, respectively;

(5) The obstacle detection and early warning module, based on the horizontal coordination position of the power line in left and right visual images ($x_{dz1}, x_{dz2}...x_{dzj}$), ($x_{dy1}, x_{dy2}...x_{dyj}$) and locations of exact matching points in left and right visual images (x_{z1}, y_{z1})~(x_{y1}, y_{y1}), (x_{z2}, y_{z2})~(x_{y2}, y_{y2})...(x_{zn}, y_{zn})~(x_{yn}, y_{yn}), uses the binocular visual principle to calculate the three-dimensional space coordinates of the matching point and the power line, and calculate the vertical distance from the matching point to the power line according to the space coordinate information;

(6) The result output and feedback module back feeds the obstacle information which threatens the power line. When the space vertical distance between the matching point and the power line is ~~higher~~ lower than the given threshold, back feed concrete information about the obstacle on the computer software interface and record the obstacle detection time and geographical location. When the space vertical distance between the matching point and the power line is ~~lower~~ higher than the given threshold, the current matching point shall not be processed and the same treatment shall be applied to all matching points in the current frame image. In addition, back feed the obstacle information about the frame images obtained in sequence to complete record of the obstacle information in the inspection process~~.~~;

wherein the characteristics are as follows:

(i) As regards Step (2), the image preprocessing module processes the left video sequence l_z and the right video sequence l_y, which first grabs the frame continuously from the left video sequence l_z and the right video sequence l_y separately at a certain interval, which requires that all frame images can be pieced together to recover the complete inspection image information and there is no information gap when adjacent frames are pieced together, so as to form the left image sequence T_z and right image sequence T_y, of which the left image sequence T_z and right image sequence T_y contain *m* frames of images, separately are $T_{z1}, T_{z2} ... T_{zm}$ and $T_{y1}, T_{y2} ... T_{ym}$; preprocess the corresponding frames of left and right visual images, and first process the first frame of left visual image T_{z1} and right visual image T_{y1}, and then process the corresponding frame of the left and right visual images in sequence until the end of processing of *m* frame of left visual image T_{zm} and *m* frame of left visual image T_{ym};

(ii) Preprocess the first frame of left visual image T_{z1} and the first image of right visual image T_{y1}, including image grayscale processing and DoG edge detection:

Grayscale processing of the said image turns color image into the single-channel grayscale image H_{z1} of the first left visual image and the single-channel grayscale image of the first right visual image H_{y1};

In the said DoG edge detection, DoG operation was defined as:

$$D = \frac{1}{2\pi}\left[\frac{1}{\sigma_1^2} e^{-\frac{(x^2+y^2)}{2\sigma_1^2}} - \frac{1}{\sigma_2^2} e^{-\frac{(x^2+y^2)}{2\sigma_2^2}}\right] * H$$
$$= [G(x, y, \sigma_1) - G(x, y, \sigma_2)] * H$$

Where, D refers to the image result after processing, σ_1 is the first constant while σ_2 is the second constant, x and y separately refer to the horizontal and vertical coordinates of the current pixel point in the image, $G(x, y, \sigma_1)$ and

$G(x, y, \sigma_2)$ separately refer to the Gaussian filter function of different parameters, H refers to the grayscale image, "*" refers to treatment of moving filter toward the whole frame of image, and the preprocessed images separately are the left visual edge image D_{z1} and the right visual edge image D_{y1};

(iii) As regards Step (3), segment the power line in the left visual edge image D_{z1} and the right visual edge image D_{y1}, $B(x)$ refers to structural element, and erosion and dilation operation of each point $A(x, y)$ in the work space E can be defined as:

Erosion: $X = E \otimes B = \{A(x, y) : B(x) \subset E\}$

Dilation: $Y = E \oplus B = \{A(x, y):\ B(y) \cap E \neq \Phi\}$

Where, $\subset$ stands for "included in", $\cap$ stands for "intersection operation" and Φ stands for null set. If the structural element $B(x)$ is used to apply erosion to the work space E, the result is the set consisting of all points of the structural element $B(x)$ included in the work space E after translation of the structural element $B(x)$. If the structural element $B(x)$ is used to apply dilation to the work space E, the result is the set consisting of all non-null set points of the structural element B included in the work space E after translation of the structural element B;

(iv) With regard to the image which has completed the loop erosion and dilation operation, measure the area and length of the area of the connected domain in the image, retain the line segment whose linear shape and area of the connected domain meet the threshold requirements, that is, the power line, remove the noise in the image, complete segmentation of the power line in the image and separately record the horizontal coordinate positions of $(x_{dz1}, x_{dz2} \ldots x_{dzj})$ and $(x_{dz1}, x_{dz2} \ldots x_{dzj})$ of the power line in the left visual edge image D_{z1} and the right visual edge image D_{y1}, where the left visual edge image D_{z1} and the right visual edge image D_{y1} separately contain j

power lines and record *j* horizontal coordinate positions; and

(v) As regards Step (4), use SURF algorithm to detect the feature points of the left visual edge image D_{z1} and the right visual edge image D_{y1}, and the specific steps are as follows:

(41) Use the rapid Hessian detector to extract the feature points, with regard to certain point in the left visual edge image D_{z1} and the right visual edge image D_{y1}, when the scale is σ, the Hessian matrix at the point shall be expressed as:

$$H(x,\sigma)=\begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix}$$

Where, the function $L_{xx}(x,\sigma)$ refers to the second partial derivative of the Gaussian function on X-coordinate and the convolution of the left visual edge image D_{z1} and the right visual edge image D_{y1} at the point (x,y), and the formula is shown below:

$$L_{xx}(x,\sigma)=D*\frac{\partial^2}{\partial x^2}g(\sigma)$$

Description of $L_{xy}(x,\sigma)$ and $L_{yy}(x,\sigma)$ is separately shown in the formula below:

$$L_{xy}(x,\sigma)=D*\frac{\partial^2}{\partial x\partial y}g(\sigma)$$

$L_{yy}(x,\sigma)=D*\frac{\partial^2}{\partial y^2}g(\sigma)$, ∂ refers to the operation to obtain the partial derivative,

Where, the scale space of SURF algorithm is divided by group, images in each group are obtained after convolution of filters of different sizes, and the filter size

will gradually increase;

(42) Set the feature point as the central point and construct a box with the size of 20σ along the principal direction of the central point, then divide the region into 16 small regions, calculate the wavelet response within each subregion ($5\sigma \times 5\sigma$) and obtain the direction of $0°$ and the vector $\sum dx$, direction of $90°$ and the vector $\sum dy$, direction of $180°$ and the vector $\sum d|x|$ and direction of $180°$ and the vector $\sum d|y|$, and then construct the 4D feature vector $v = (\sum dx, \sum |dx|, \sum dy, \sum |dy|)$ to express each subregion, and finally form the 64D descriptor of the point;

Assuming that the size of the left visual edge image D_{z1} and the right visual edge image D_{y1} are both $M \times N$, and the images are placed horizontally at the same coordinate axis to form the image with the size of $M \times 2N$, and the set of feature points of the left visual edge image D_{z1} and the right visual edge image D_{y1} detected by the SURF method can be expressed as:

$$Pos1 = \{(x_1^{'}, y_1^{'}), (x_2^{'}, y_2^{'}), ..., (x_p^{'}, y_p^{'})\}$$

$$Pos2 = \{(x_1, y_1), (x_2, y_2), ..., (x_q, y_q)\},$$

Where, p and q separately refer to the number of feature points of the left visual edge image D_{z1} and the right visual edge image D_{y1};

Feature point matching method includes the following steps:

a) With regard to each point i in the feature point set $Pos1$ of the left visual edge image, calculate the Euclidean distance of it from all points in the feature point set $Pos2$ of the right visual edge image, and select the corresponding point of the

minimum Euclidean distance as the rough matching point of Point i;

b) Calculate the Euclidean distance of all rough matching point pair, sort the matching point pairs by the Euclidean distance in ascending order and delete the point pair with multiple points towards one point, at this time the feature points in the left visual edge image D_{z1} and the right visual edge image D_{y1} can be separately expressed as the feature point set $Pos1'$ of the fixed left visual edge image and the feature point set $Pos2'$ of the fixed right visual edge image;

c) Select the first K_1 pairs of matching points in the feature point set $Pos1'$ of the fixed left visual edge image and the feature point set $Pos2'$ of the fixed right visual edge image to express as $Pos_K_1 = \{\{(x_1', y_1'), (x_1, y_1)\}, \{(x_2', y_2'), (x_2, y_2)\}, ..., \{(x_n', y_n'), (x_n, y_n)\}\}$, which is called Set 1;

Select the first K_2 pairs of matching points in the feature point set $Pos1'$ of the fixed left visual edge image and the feature point set $Pos2'$ of the fixed right visual edge image to express as Pos_K_2, where $Pos_K_2 = \{\{(x_1', y_1'), (x_1, y_1)\}, \{(x_2', y_2'), (x_2, y_2)\}, ..., \{(x_{K_2}', y_{K_2}'), (x_{K_2}, y_{K_2})\}\}$, is called as Set 2, where $K_1 < K_2$;

d) With regard to all the point pairs in Set 2, calculate the slope between two points as shown in the formula below, and round it to form Slope Set k: $k = \{k_1, k_2, ..., k_w\}$,

$$k_w = \frac{y_w}{x_w - x_w'}, 1 \le w \le K_2;$$

e) Calculate the frequency of each slope in Slope Set k, screen the slope with the

frequency of greater than and equal to 2 to form a new set $k_new = \{k_1, k_2, ..., k_t\}$, where t refers to the total quantity of feature points obtained; if the frequency of each slop in Slope Set k is 1, select the slope of the first $2K_2/3$ pairs of points to form a new set k_new;

f) Traverse and calculate the slope of all point pairs in the feature point set $Pos1'$ of the fixed left visual edge image and the feature point set $Pos2'$ of the fixed right visual edge image, and screen all the point pairs with the slope falling between the interval of $[k_t - 0.5, k_t + 0.5]$ to form the point pair set

$$Pos_K_3 = \{\{(x_{z1}, y_{z1}), (x_{y1}, y_{y1})\}, \{(x_{z2}, y_{z2}), (x_{y2}, y_{y2})\}, ..., \{(x_{zn}, y_{zn}), (x_{yn}, y_{yn})\}\}$$, where $k_t \in k_new$.

2. According to the UAV inspection method of the power line based on human visual system as stated in Claim 1, the characteristics are as follows: In Step (1) as described, the binocular visual measuring equipment as described refers to two video image capture equipment with the same specifications and parameters, which are arranged in the left and right visual forms separately and acquire video image information at a fixed angle of view;

When the UAV armed with binocular visual measuring equipment is used for power line inspection, wireless remote control shall be applied to the UAV, so that the UAV can have uniform rectilinear movements on the path in parallel with the power line and above the power line, and the distribution direction of the power line in the capture video image sequence is in parallel with the flight direction of the UAV;

The power line video image captured includes the left video sequence l_z and right video sequence l_y, and stores the data of the left video sequence l_z and the right video sequence l_y in a storage device installed on the UAV, which shall be transmitted to the

image preprocessing module of the software system via the mobile wireless network.

3. According to the UAV inspection method of the power line based on human visual system as stated in Claim 1, the characteristics are as follows: In Step (2) as described, the image preprocessing module processes the left video sequence l_z and the right video sequence l_y, which first grabs the frame continuously from the left video sequence l_z and the right video sequence l_y separately at a certain interval, which requires that all frame images can be pieced together to recover the complete inspection image information and there is no information gap when adjacent frames are pieced together, so as to form the left image sequence T_z and right image sequence T_y, of which the left image sequence T_z and right image sequence T_y contain m frames of images, separately are $T_{z1}, T_{z2} \ldots T_{zm}$ and $T_{y1}, T_{y2} \ldots T_{ym}$; preprocess the corresponding frames of left and right visual images, and first process the first frame of left visual image T_{z1} and right visual image T_{y1}, and then process the corresponding frame of the left and right visual images in sequence until the end of processing of m frame of left visual image T_{zm} and m frame of left visual image T_{ym}.

4. According to the UAV inspection method of the power line based on human visual system as stated in Claim 3, the characteristics are as follows: Preprocess the first frame of left visual image T_{z1} and the first image of right visual image T_{y1}, including image grayscale processing and DoG edge detection:

Grayscale processing of the said image turns color image into the single-channel grayscale image H_{z1} of the first left visual image and the single-channel grayscale image of the first right visual image H_{y1};

In the said DoG edge detection, DoG operation was defined as:

$$D = \frac{1}{2\pi}\left[\frac{1}{\sigma_1^2}e^{-\frac{(x^2+y^2)}{2\sigma_1^2}} - \frac{1}{\sigma_2^2}e^{-\frac{(x^2+y^2)}{2\sigma_2^2}}\right] * H$$
$$= [G(x,y,\sigma_1) - G(x,y,\sigma_2)] * H$$

~~Where, D refers to the image result after processing, σ_1 is the first constant while σ_2 is the second constant, x and y separately refer to the horizontal and vertical coordinates of the current pixel point in the image, $G(x,y,\sigma_1)$ and $G(x,y,\sigma_2)$ separately refer to the Gaussian filter function of different parameters, H refers to the grayscale image, "*" refers to treatment of moving filter toward the whole frame of image, and the preprocessed images separately are the left visual edge image D_{z1} and the right visual edge image D_{y1}.~~

~~5. According to the UAV inspection method of the power line based on human visual system as stated in Claim 4, the characteristics are as follows: In the said Step (3), segment the power line in the left visual edge image D_{z1} and the right visual edge image D_{y1}, $B(x)$ refers to structural element, and erosion and dilation operation of each point $A(x,y)$ in the work space E can be defined as:~~

~~Erosion: $X = E \otimes B = \{A(x,y) : B(x) \subset E\}$~~

~~Dilation: $Y = E \oplus B = \{A(x,y).\ B(y) \cap E \neq \Phi\}$~~

~~Where, $\subset$ stands for "included in", $\cap$ stands for "intersection operation" and Φ stands for null set. If the structural element $B(x)$ is used to apply erosion to the work space E, the result is the set consisting of all points of the structural element $B(x)$ included in the work space E after translation of the structural element $B(x)$. If the structural element $B(x)$ is used to apply dilation to the work space E, the result is the set consisting of all non-null set points of the structural element B included in the work space E after translation of the structural element B.~~

3~~6~~. According to the UAV inspection method of the power line based on human visual system as stated in Claim ~~5~~1, paragraph (iii), the characteristics are as follows: Select the linear structure factor with the length of 3 pixel points and the angle of 90° , apply erosion and dilation operation to the left visual edge image D_{z1} and the right visual edge image D_{y1}, during which one loop operation includes two erosion operations and one dilation operation, which lasts for 20 times.

~~7. According to the UAV inspection method of the power line based on human visual system as stated in Claim 5, the characteristics are as follows: With regard to the image which has completed the loop erosion and dilation operation, measure the area and length of the area of the connected domain in the image, retain the line segment whose linear shape and area of the connected domain meet the threshold requirements, that is, the power line, remove the noise in the image, complete segmentation of the power line in the image and separately record the horizontal coordinate positions of $(x_{dz1}, x_{dz2} \ldots x_{dzj})$ and $(x_{dz1}, x_{dz2} \ldots x_{dzj})$ of the power line in the left visual edge image D_{z1} and the right visual edge image D_{y1}, where the left visual edge image D_{z1} and the right visual edge image D_{y1} separately contain j power lines and record j horizontal coordinate positions.~~

~~8. According to the UAV inspection method of the power line based on human visual system as stated in Claim 5, the characteristics are as follows: In the said Step (4), use SURF algorithm to detect the feature points of the left visual edge image D_{z1} and the right visual edge image D_{y1}, and the specific steps are as follows:~~

~~(41) Use the rapid Hessian detector to extract the feature points, with regard to certain point in the left visual edge image D_{z1} and the right visual edge image D_{y1}, when the scale is σ, the Hessian matrix at the point shall be expressed as:~~

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix}$$

~~Where, the function $L_{xx}(x,\sigma)$ refers to the second partial derivative of the Gaussian function on X-coordinate and the convolution of the left visual edge image D_{z1} and the right visual edge image D_{y1} at the point (x,y), and the formula is shown below:~~

$$L_{xx}(x,\sigma)=D*\frac{\partial^2}{\partial x^2}g(\sigma)$$

Description of $L_{xy}(x,\sigma)$ and $L_{yy}(x,\sigma)$ is separately shown in the formula below:

$$L_{xy}(x,\sigma)=D*\frac{\partial^2}{\partial x\partial y}g(\sigma)$$

$L_{yy}(x,\sigma)=D*\frac{\partial^2}{\partial y^2}g(\sigma)$, ∂ refers to the operation to obtain the partial derivative,

Where, the scale space of SURF algorithm is divided by group, images in each group are obtained after convolution of filters of different sizes, and the filter size will gradually increase;

(42) Set the feature point as the central point and construct a box with the size of 20σ along the principal direction of the central point, then divide the region into 16 small regions, calculate the wavelet response within each subregion ($5\sigma\times5\sigma$) and obtain the direction of $0°$ and the vector $\sum dx$, direction of $90°$ and the vector $\sum dy$, direction of $180°$ and the vector $\sum d|x|$ and direction of $180°$ and the vector $\sum d|y|$, and then construct the 4D feature vector $v=(\sum dx,\sum|dx|,\sum dy,\sum|dy|)$ to express each subregion, and finally form the 64D descriptor of the point;

Assuming that the size of the left visual edge image D_{z1} and the right visual edge image D_{y1} are both $M\times N$, and the images are placed horizontally at the same coordinate axis to form the image with the size of $M\times2N$, and the set of feature points of the left visual edge image D_{z1} and the right visual edge image D_{y1}

detected by the SURF method can be expressed as:

$$Pos1 = \{(x_1^{'}, y_1^{'}), (x_2^{'}, y_2^{'}), \ldots, (x_p^{'}, y_p^{'})\}$$

$$Pos2 = \{(x_1, y_1), (x_2, y_2), \ldots, (x_q, y_q)\},$$

Where, p and q separately refer to the number of feature points of the left visual edge image D_{z1} and the right visual edge image D_{y1};

Feature point matching method includes the following steps:

a) With regard to each point i in the feature point set $Pos1$ of the left visual edge image, calculate the Euclidean distance of it from all points in the feature point set $Pos2$ of the right visual edge image, and select the corresponding point of the minimum Euclidean distance as the rough matching point of Point i;

b) Calculate the Euclidean distance of all rough matching point pair, sort the matching point pairs by the Euclidean distance in ascending order and delete the point pair with multiple points towards one point, at this time the feature points in the left visual edge image D_{z1} and the right visual edge image D_{y1} can be separately expressed as the feature point set $Pos1^{'}$ of the fixed left visual edge image and the feature point set $Pos2^{'}$ of the fixed right visual edge image;

c) Select the first K_1 pairs of matching points in the feature point set $Pos1^{'}$ of the fixed left visual edge image and the feature point set $Pos2^{'}$ of the fixed right visual edge image to express as

$Pos_K_1 = \{\{(x_1^{'}, y_1^{'}), (x_1, y_1)\}, \{(x_2^{'}, y_2^{'}), (x_2, y_2)\}, \ldots, \{(x_n^{'}, y_n^{'}), (x_n, y_n)\}\}$, which is called Set 1;

Select the first K_2 pairs of matching points in the feature point set $Pos1^{'}$ of the fixed

~~left visual edge image and the feature point set $Pos2'$ of the fixed right visual edge image to express as Pos_K_2, where $Pos_K_2 = \{\{(x_1', y_1'), (x_1, y_1)\}, \{(x_2', y_2'), (x_2, y_2)\}, \ldots, \{(x_{K_2}', y_{K_2}'), (x_{K_2}, y_{K_2})\}\}$, is called as Set 2, where $K_1 < K_2$;~~

~~d) With regard to all the point pairs in Set 2, calculate the slope between two points as shown in the formula below, and round it to form Slope Set $k: k = \{k_1, k_2, \ldots, k_w\}$,~~

$$k_w = \frac{y_w}{x_w - x_w'}, 1 \le w \le K_2$$

~~;~~

~~e) Calculate the frequency of each slope in Slope Set k, screen the slope with the frequency of greater than and equal to 2 to form a new set $k_new = \{k_1, k_2, \ldots, k_t\}$, where t refers to the total quantity of feature points obtained; if the frequency of each slop in Slope Set k is 1, select the slope of the first $2K_2/3$ pairs of points to form a new set k_new;~~

~~f) Traverse and calculate the slope of all point pairs in the feature point set $Pos1'$ of the fixed left visual edge image and the feature point set $Pos2'$ of the fixed right visual edge image, and screen all the point pairs with the slope falling between the interval of $[k_t - 0.5, k_t + 0.5]$ to form the point pair set~~

~~$Pos_K_3 = \{\{(x_{z1}, y_{z1}), (x_{y1}, y_{y1})\}, \{(x_{z2}, y_{z2}), (x_{y2}, y_{y2})\}, \ldots, \{(x_{zn}, y_{zn}), (x_{yn}, y_{yn})\}\}$, where $k_t \in k_new$.~~

~~4~~9. According to the UAV inspection method of the power line based on human

visual system as stated in Claim ~~8~~1, paragraph (v), the characteristics are as follows: In the said Step (41), SURF algorithm re-specifies the only direction for each interest point according to the information about the pixel point around the feature points, and the specific steps are as follows:

a) Take feature points as center to calculate the horizontal and vertical responses of Harr wavelet in the circular field with a radius of 6σ, of which the sampling step length is σ and the wavelet size is 4σ;

b) Take feature points as center to apply Gaussian weighting to wavelet response, so that the weight value close to the feature point is large and that away from the feature point is small, and the new horizontal and vertical responses are obtained;

c) Finally, use a fan-shaped window with an angle of 60° to traverse the entire circle until the total response within the fan-shaped window is the strongest, at this time the direction within the fan-shaped window is the principal direction of the interest point.

5~~10~~. According to the UAV inspection method of the power line based on human visual system as stated in Claim ~~8~~1, paragraph (v), the characteristics are as follows: In the said Step (5), the parallax range between binocular visual cameras is b, the known focal length of camera is f and parallax d is defined as the position difference of certain point between the corresponding points in two images;

$$d = (x_{zn} - x_{yn})$$

Where, x_{zn}、x_{yn} separately refer to the horizontal coordinates of the matching

point in the left visual image and right visual image, so as to calculate the space coordinates of certain point *P* in the left camera coordinate system as:

$$\begin{cases} x^c = \dfrac{b \cdot x_{zn}}{d} \\ y^c = \dfrac{b \cdot y_{zn}}{d} \\ z^c = \dfrac{b \cdot f}{d} \end{cases}$$

Where, (x^c, y^c, z^c) is the position information about the current matching point in the space coordinate and y_{zn} is the vertical coordinate of the matching point in the left and right visual images. Calculate the coordinates of all matching points in the left camera coordinate system according to the formula above, and then calculate the space coordinate of the point on the power line with the minimum Euclidean distance from the space of the matching point according to the formula above and the obtained horizontal coordinate positions of the power line $(x_{dz1}, x_{dz2}...x_{dzn})$ and $(x_{dy1}, x_{dy2}...x_{dyn})$; where the points defined on the 2D coordinate system with the same vertical coordinate have the minimum Euclidean distance from the space of the matching point, then directly give the vertical coordinate of the matching point to the corresponding point on the power line to form a corresponding point of the power line $Pos_D = \{\{(x_{dz1}, y_{z1}), (x_{dy1}, y_{y1})\}, \{(x_{dz2}, y_{z2}), (x_{dy2}, y_{y2})\}, ..., \{(x_{dzn}, y_{zn}), (x_{dyn}, y_{yn})\}\}$ with the point pair set $Pos_K_3 = \{\{(x_{z1}, y_{z1}), (x_{y1}, y_{y1})\}, \{(x_{z2}, y_{z2}), (x_{y2}, y_{y2})\}, ..., \{(x_{zn}, y_{zn}), (x_{yn}, y_{yn})\}\}$, thus to calculate the space coordinate (x^d, y^d, z^d) of certain point *D* on the power line in the left camera coordinate system.

6~~11~~. According to the UAV inspection method of the power line based on human visual system as stated in Claim ~~10~~5, the characteristics are as follows: In the said Step (6), calculate the Euclidean distance *J* of the matching point from the space of the power line after the space coordinates of Point *P* and Point *D* are obtained, and *J* is defined as:

$$J = \sqrt{(x^c - x^d)^2 + (y^c - y^d)^2 + (z^c - z^d)^2} \quad .$$