STUDY ON THREE-DIMENSIONAL RECONSTRUCTION OF GREEN PLANTS LANDSCAPE ON BOTH SIDES OF HIGHWAY

Yang Yao

Chongqing Jiaotong University, College of architecture and urban planning, Chongqing, China
E-mail: Yaoyao121121@tom.com

ABSTRACT: The current three-dimensional landscape modeling cannot reflect the ups and downs of green plants on both sides of highway, and the matching errors of feature points are large and reconstruction quality is low. A three-dimensional landscape reconstruction method of green plants on both sides of highway is proposed. The three-dimensional information of the green plants on both sides of the highway is calculated by the obtained image sequences. After the more complete edge image pairs are extracted, the SIF algorithm is used to extract the corner points of the image, and the SIFT feature subspace is constructed to obtain the feature descriptor for initial matching, the matching cost function based on adaptive weights is constructed, and the matching point with the lowest cost function is found as the matching point of the point. The point clouds of two images are spliced together to complete the three-dimensional landscape reconstruction by using ICP algorithm. Experimental results show that this method can effectively improve the matching accuracy of feature points and the reconstruction quality.

KEYWORDS: Both Sides of Highway, Green Plants, Three-Dimensional Reconstruction of Landscape

1 INTRODUCTION

In the driving process on the highway, the information from outside world is usually introduced by the vision. Thus, changing the visual field of the driver is very important, and long-time driving at high speed on the highway will make mental and visual fatigue, and detrimental driving (Ohtake et al., 2015; Ficker & Martisek, 2016). Three-dimensional (3D) reconstruction of the landscape of green plants on both sides of highway and the rational planning of landscapes of green plants on both sides of highway can greatly enhance the realism and immersion of the plant landscape and attract the attention of drivers, so as to adjust the vision and eliminate mental fatigue (Dahl et al., 2015; Wakui et al., 2015). However, due to the complexity and morphological diversity of green plants on both sides of highway, the three-dimensional landscape reconstruction and the rational planning of landscapes of green plants in real environment is a challenging issue in computer graphics and virtual visualization (Zhen, 2016).

Currently, the effective green building landscape modeling methods for both sides of highway are as follows: The method proposed by Wu et al. (2016) is a 3D reconstruction method based on on-board LiDAR data. In the process of reconstruction, it is hard to take into account the local feature information, which will lead to discontinuity of outline edge of green plant landscape and low reconstruction quality. Zhao (Zhao, Sun & Chen, 2016) reported a 3D image reconstruction method using acceleration of a micro-inertial sensor. This method only takes into account the influence of camera shooting. There are too many pseudo-feature points in the feature points after the actual edge extraction, and the quality of the 3D reconstruction of the green plant landscape is low. Therefore, a 3D reconstruction method of green plant landscape on both sides of highway is proposed.

2 CAMERA MODEL AND CALIBRATION METHOD

Camera calibration is an indispensable step in getting 3D information from 2D images. Camera parameters are always relative to some geometrical imaging model, which is a simplification of the optical imaging process (Konold & Jimenez, 2015). World coordinate system is converted to camera coordinate system, because the world coordinate system is the coordinate system we take in the natural environment, we can regard it as the camera coordinate system of the first highway green map, and the conversion can be expressed as follows:
The epipolar geometric relationship between two images is a fundamental concept in computer vision. It is defined by epipolar lines and epipolar points. The epipolar line of a point in one image is the line in the other image that contains the matching point. The epipolar points are the corresponding points in the two images.

The epipolar geometric relationship can be described by the following equations:

\[
\begin{bmatrix}
X_e \\
Y_e \\
Z_e
\end{bmatrix} = \begin{bmatrix}
R & t \\
0^T & 1
\end{bmatrix} \begin{bmatrix}
X_m \\
Y_m \\
Z_m
\end{bmatrix}
\]

Where \( R \) is the rotation matrix, \( t \) is the translation vector, and \( 0^T \) is the column vector with element 0. Camera imaging meets the pinhole imaging model, and the conversion process in line with the perspective projection, which can be expressed with homogeneous coordinates as follows:

\[
\begin{bmatrix}
x \\
y \\
1
\end{bmatrix} = \begin{bmatrix}
f & 0 & 0 & 0 \\
0 & f & 0 & 0 \\
0 & 0 & 1 & 0 \\
1
\end{bmatrix} \begin{bmatrix}
X_e \\
Y_e \\
Z_e
\end{bmatrix}
\]

(2)

Where \((X_e, Y_e, Z_e)\) is the coordinate of a space point \( P \) in the optical coordinate system, \((x, y)\) is the coordinate of the point \( P \) in the image coordinate system, and \( f \) is the focal length of the camera. Conversion from image coordinate system to pixel coordinate system is:

\[
\begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix}
1/dx & 0 & u_0 \\
0 & 1/dy & v_0
\end{bmatrix} \begin{bmatrix}
x \\
y \\
1
\end{bmatrix}
\]

(3)

Where \((u_0, v_0)\) is the coordinate of the origin of the image coordinate system in the pixel coordinate system, and \((dx, dy)\) is the distance between adjacent pixels in the \( X \) and \( Y \) directions of the pixel coordinate system, respectively.

The point \( P \) on the first image of the green plant on both sides of highway is the image point of all the space points on the radial \( t \) that passes through this point and the center of the camera lens. The point \( P' \) on the second image of the green plant on both sides of highway corresponding to point \( P \), is on the epipolar line \( L \) of point \( P \). If the epipolar equation is \( px + qy + r = 0 \), \( L \) can be expressed as \((p, q, r)^T\). Therefore, all spatial points on \( L \) should be on line \( L \) in the second image. The center of the lens \((0,0,1)^T\) and the infinity point \([wA^{-1}(u,v,1)^T, 0]\) of the first image of green plants on both sides of highway are at \( l \), and pixels of all the points on \( l \) are on the first image of green plants on both sides of highway. The pixels of the two points on the second green plant image are \(-A^RT\) and \(wA^RA^{-1}(u,v,1)^T\), respectively, then the epipolar line \( L \) of point \( P \) is:

\[
(p, q, r)^T = wA^RT \times A^RA^{-1}(u,v,1)^T
\]

(4)

If \( K \) is a square matrix, let \( K' = (K_{ij}) \), \( K_{ij} \) is the algebraic remainder of the element \( K_{ij} \) of the matrix \( K \). If \( K \) is a non-singular matrix, then \( K' = \det(K)(k^T)^{-1} \).

3 EPIPOLAR GEOMETRIC BASIS AND BASIS MATRIX SOLUTION METHOD

For the same scene, the sequences of images acquired from different viewpoints satisfy certain relations, which exist in two or more green plant images. There is an epipolar geometric relationship between the two images, which is the theoretical basis for the realization of 3D reconstruction.

In the figure, \( C \) and \( C' \) are the optical centers of the two cameras respectively. The intersections \( e \) and \( e' \) of \( C \) and \( C' \) with the left and right images are called epipolar points, and the corresponding coordinates are \( \theta = (e, e, 1)^T \) and \( \theta' = (e', e', 1)^T \). \( M \) is one point in the world coordinate system. That is, \( m \) and \( m' \) are the image points of \( M \) in left and right images. The plane \( \pi \) defined by \( M \) and \( C \) and \( C' \) is called the epipolar plane. The line \( l_m \), defined by points \( m \) and \( e \) is called the epipolar line of the point \( m \) in the right image, \( l_m = (n, \theta) \).

The line \( l_m \) is defined by points \( m \) and \( e \) and \( e' \) is called the epipolar line of the point \( m \) and \( e \) in the left image, \( l_m^L = (n, \theta) \). For the point \( m \) in the left image, its corresponding point in the left image must lie on the epipolar line \( l_m^L \) of point \( m \). For the point \( m' \) in the right image, its matching point must be located on the epipolar line \( l_m \) of point \( m' \). The epipolar relationship can be described as follows:

\[
\begin{align*}
l_m^L \cap K' &= (n, \theta) \\
l_m \cap F_m &= (n, \theta)
\end{align*}
\]

(5)

\[
\begin{align*}
l_m^L \cap F_m &= (n, \theta) \\
l_m \cap F_m &= (n, \theta)
\end{align*}
\]

(6)

\[
\begin{align*}
l_m^L \cap F_m &= (n, \theta) \\
l_m \cap F_m &= (n, \theta)
\end{align*}
\]

(7)
In the basic matrix solution process, the two green plant images are processed as follows:

(1) Coordinate transformation: find the center of gravity of the green plant image and make a translation transformation to the image coordinate system, so that the original image coordinate system with the upper left corner as the origin translates to the image with the origin of all the green plant image points as the origin coordinate System.

(2) Scale Transformation: make a scale transformation on the coordinates of the green plant image so that the average distance from each point to the origin is \( \sqrt{2} \). The coordinate of matching points in the new coordinate system is obtained and regarded as the input data to calculate the basic matrix, the process is as follows:

Let two independent green plant image coordinate transformations be \( T \) and \( T' \) respectively, then the new image coordinates are \( m_n = Tn \) and \( m'_n = T'n \). The base matrix \( F_n \) is calculated using the eight-point algorithm with the new matching points \( m_n \) and \( m'_n \). The original base matrix is

\[
F = T^T F_n T.
\]

4 PLAN TANSWORLD 3D IMAGE FEATURE EXTRACTION

4.1 Edge detection of 3D image

The optimized Canny operator (Li, Zhao & Fu, 2015) is used to obtain the complete edge information of green plants. SIFT (Hu et al., 2015) is used to extract the corner points of green plants on both sides of highway, and remove false feature points, so as to get the accurate information needed in the image matching of green plants on both sides of highway. The gradient value with the most pixels in the gradient histogram is defined as the value of the pixel gradient \( H_{\text{max}} \), the gradient of all pixels in sub-image is the variance of the most significant pixel gradient \( H_{\text{max}} \), known as the pixel value of the gradient \( \sigma_{\text{max}} \), calculated as follows:

\[
\sigma_{\text{max}} = \sqrt{\frac{\sum_{i=0}^{N}(H_i - H_{\text{max}})^2}{N}}.
\]

Where, \( n \) is the maximum value of gradient with non-0 pixels, \( N \) is the total number of pixels. When there is only a single peak in the gradient histogram, the gradient values of the pixels are all concentrated in the vicinity of the maximum value gradient \( H_{\text{max}} \), and the variance \( \sigma_{\text{max}} \) of the pixel maximum gradient is very small. When the gradient histogram not only has a single peak of the non-edge pixel gradient, but also the distribution of the edge pixel gradient, the edge pixel gradient is distributed at a relatively far distance from the most value gradient \( H_{\text{max}} \), and the pixel maximum gradient variance \( \sigma_{\text{max}} \) is large. Therefore, threshold can be set to judge.

4.2 Feature point extraction

SIFT algorithm is used to detect the feature of green plant landscape image space on both sides of highway to determine the location of the key points and the scale from which they are located. Based on the main direction of the neighborhood gradient of the key points, the independence of the operator from the scale and the direction are achieved. The scale space of a green plant landscape image on both sides of highway is defined as:

\[
L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)
\]

Where, \( G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \)

Where, \( e^\theta \) represents the scale space coefficient. In order to effectively detect the location of stable feature points, a DOG pyramid is used in the scale space. Fitting function is a second order Taylor expansion of DOG operator:

\[
D(\vec{\gamma}) = D + \frac{\partial D}{\partial \vec{\gamma}} \vec{\gamma} - \frac{1}{2} \frac{\partial^2 D}{\partial \vec{\gamma} \partial \vec{\gamma}^T} \vec{\gamma} \vec{\gamma}^T
\]

If the value of \( D(\vec{\gamma}) \) at an extreme point is less than a certain closed value, the point is unstable and will be removed. Since the principal curvature of \( D \) is proportional to the eigenvalue of the matrix \( H \), the eigenvalues are not specified and the ratio is calculated. Let \( \alpha \) be the largest amplitude feature, \( \beta \) be the second smallest, \( r = \alpha / \beta \), then:

\[
\text{ratio} = \frac{\text{Tr}(H)^2}{\text{Det}(H)} = \frac{(\alpha + \beta)^2}{\alpha \beta} = \frac{(r+1)^2}{r}
\]

In the feature point, if \( \text{ratio} \) is meet, then retain otherwise remove. Thus, removing the boundary point, the key points of the green plant landscape images on both sides of highway have been detected.
5 GREEN PLANTS LANDSCAPE IMAGES MATCHING

5.1 SIFT feature subspaces

The SIFT feature description is obtained using a histogram of gradient direction histograms of a 16 × 16 size window around the feature point. Supposing a certain feature point \( P' \), \( a \), \( b \) and \( c \) describe the shape of \( P \) under different image scales \( \delta_a \), \( \delta_b \) and \( \delta_c \), respectively. Assuming that the scale space of the image is \( \delta = \{ \delta_L, \delta, \delta_s \} \), and the feature vector of \( P \) calculated at different scales is \( v_i \), the feature subspace is generated by \( V = span(v_i, L, v) \), where a set of basis \( V \) is \( v_i, L, v_a \). Thus \( \dim(V) = k \), then the description of \( P \) can be obtained by a linear combination of the basis vectors in \( V \), defined as:

\[
Det(P) = w_1^j v_{ij} + w_2 v_{12} + L + w_k v_{id}
\]

Where, \( w_j \) is the weight of \( v_i \), and in order to maintain the unit of vector, we need to make \( \sum_{j=1}^{k} w_j = 1 \). The choice of the weight value depends on the contribution of the scale image of \( v_i \) to the \( P \)-point.

5.2 Initial Match

After extracting and describing the feature points, the KD-Tree algorithm is used to calculate the Euclidean distances to obtain the initial matching point pairs. Assuming that \( X \) and \( Y \) are two SIFT point sets of matching images. The corresponding initial matching point pairs are obtained by the following algorithm:

1. For a certain feature point \( x_i \) in \( X \), calculate the distance \( dist(x_i, y_j) \) between each \( y_j \) and \( x_i \) in \( Y \).
2. Select the two of the minimum distance \( D_n \) and \( D_s \), calculate the ratio \( \text{ratio} = D_n / D_s \).
3. Select the ratio threshold \( T_{\text{ratio}} \), if \( \text{ratio} < T_{\text{ratio}} \), then consider that \( x_i \) and \( y_j \) are matched, add it to the corresponding points set \( Z \), otherwise stop and proceed to the next calculation of \( x_{i+1} \).
4. Repeat the above steps until all the feature points in \( X \) are calculated.

5.3 Local graphics approximation

In order to obtain an accurate transformation \( T \), it is necessary to select the point-pair set with the highest similarity in the local area as the sample pair for calculating the transformation \( T \):

\[
T_o = \arg \max S(T(I_k), I_m)
\]

Where, \( S(T(I_k), I_m) \) represents the similarity of the image area \( I_m \) with the image area \( I_k \) after passing through the transformation \( T \). Choose a non-negative depth function \( f_0 \) and a weight function \( w_0 \), compute the mean and covariance matrix by iterative operation of the Huber algorithm to obtain the correlation coefficient \( \rho \), and select \( P \) with the largest correlation coefficient to obtain the best transformation \( T \). The sample correlation coefficient method directly calculates the sample mean, variance and covariance matrix through the sample, so as to obtain the correlation coefficient:

\[
\hat{\rho}(X,Y) = \hat{\beta}_s(X,Y) / \sqrt{\hat{\sigma}_x^2 \hat{\sigma}_y^2}
\]

Where, \( \hat{\beta}_s(X,Y) \), \( \hat{\sigma}_x^2 \) and \( \hat{\sigma}_y^2 \) represent the sample mean, sample variance and sample covariance, respectively.

Because of the SIFT feature, green plant landscape images on both sides of the highway are initially matched, there are usually mismatches. Four pairs of matching points can be calculated the centroid of each of the resulting geometries separately, then describe the degree of matching between the two graphs by the difference in centroid. The difference was defined as:

\[
S(c_x, c_y) = f_0(\| c_x - c_y \|)
\]

Where, \( c_x \) and \( c_y \) are the centroids of two green plant landscape images and \( f_0 \) is a monotone non-negative function.

6 THREE-DIMENSIONAL IMAGE MOSAIC POINT CLOUD OF GREEN PLANT LANDSCAPE

ICP Algorithm was applied to point cloud matching problem of three-dimensional targets in different plant landscape spaces. Suppose two pairs of points \( P' = \{ p_1, p_2, K, p_n \} \) and \( Q = \{ q_1, q_2, K, q_n \} \) in space, look for a rigid transformation \( T \) that the following is the smallest:

\[
f(T) = \sum_{i=1}^{n} \| q_i - T(p_i) \|^2
\]
The goal of the ICP is to find the rotation transformation $R$ and the translation transformation $t$ between the point set and the point set that the transformations $T(p_i) = R_{ik} + t$ and $q_i$ are closest in distance. Suppose the set of calculated points $P'$ in the $k$th iteration is transformed into $P' = \{p'_1, p'_2, ..., p'_n\}$ to calculate the transformation matrix $T_k$, and the original transformation $T_{k-1}$ is updated until the distance between the data is less than the given threshold value $th$. According to the ICP algorithm, the green plant landscape three-dimensional image space point cloud obtained from the disparity map can be stitched together to define a new adaptive cost function, which can be expressed as follows:

$$E(p, q) = \sum_{q \in \mathcal{S}} w(p, q)w(q, \bar{q})\Psi(q, \bar{q})l(w(p, q)w(q, \bar{q}))$$

(18)

Where, $p_d$ and $q_d$ represent the corresponding points in the left and right image windows of the green plant landscape, respectively, and the corresponding parallax is $d$, which is used to describe the spatial pixel difference of the corresponding point in the window.

7 EXPERIMENTAL RESULTS

In order to prove the validity of the three-dimensional reconstruction method of green plants on both sides of highway proposed in this paper, a simulation experiment was carried out by using C++, which is testing on a desktop equipped with Intel Core TM2 2.8GHz CPU and 4GB memory. The establishment of three-dimensional landscape was using OpenGL or Visual LISP, respectively. In order to ensure that the images of green plants obtained on both sides of the highway are not blurred and the exposure time of the camera is under 1/500 second, experiments on multiple sets of data are performed to verify the stability and validity of the proposed three-dimensional reconstruction method.

Image toolbox was using on Matlab7.0 platform to achieve image edge extraction results of different methods. Take the test image 1 of Figure 3 as an example, the extraction results are as follows:

(A) Document edge extraction of Wu’s method.

(B) Document edge extraction of Zhao’s method.

(C) Document edge extraction of the proposed method.

Fig.2. Document edge extraction of different methods

Analysis of Figure 2 showed that there are more pseudo-corners of the green plant patterns on both sides of the highway extracted from the method in Figure 2 (B). Figure 2 (C) showed that the edge of the image extracted by this method further obtains the complete green plant edge map, and the outline of the cloth set on the plant leaves is also clearly visible.

the matching method proposed in this paper was tested with Figure 3 and Figure 4. The test image 2 was rotated and translated with respect to the test image 1, and the light was also different. The corner points of the green plants on both sides of the two highways are sorted, the most obvious characteristics of each map 50 points are taken through the proposed method to find the two maps in the 24 pairs of matching points and all matching correctly.

Fig.3. Test image 1
Through a large number of experiments, it showed that the proposed matching method is more than 97.52% for the sequence matching of green plants on both sides of the highway, and is invariant to rotation, translation and illumination.

Table 1. Matching experimental results

<table>
<thead>
<tr>
<th>Experimental images</th>
<th>The proposed method</th>
<th>Zhao method</th>
<th>Wu method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>100%</td>
<td>85.47%</td>
<td>92.25%</td>
</tr>
<tr>
<td>Group 2</td>
<td>97.52%</td>
<td>59.25%</td>
<td>84.58%</td>
</tr>
<tr>
<td>Group 3</td>
<td>98.58%</td>
<td>65.25%</td>
<td>79.58%</td>
</tr>
</tbody>
</table>

Analysis of Table 1 showed that the method of Zhao has a larger matching error rate for the key points with the closest neighborhood information. However, the method in this paper has high accuracy in key point matching process, which can effectively eliminate mismatch. It can be seen that the proposed method still needs to further improve the time complexity if it is to be used for processing real-time information. Therefore, the combination of stereo vision, depth camera, structured light and other technologies will make up for the shortcomings of using a single technology and improve the accuracy of acquiring 2D plant information of green plants on both sides of the highway and broaden the application range.

8 CONCLUSIONS

This paper mainly studies the acquisition of green plant data and the construction of three-dimensional virtual landscape on both sides of the highway by computer vision. For the three-dimensional data acquisition, first calibrates the camera for internal parameters, with improved Canny algorithm to extract more complete edge image pairs based on the use of SIF algorithm for image corner extraction. The method of stereo matching based on parallel distribution estimation algorithm is combined with the popular evolutionary algorithm and the idea of adaptive weight, and some improvements are made for the dense matching of images. Three-dimensional point cloud was obtained by stereo matching, the spatial point cloud of two frames of images were obtained by using the ICP algorithm, thereby the three-dimensional reconstruction of the landscape of green plants on both sides of the highway was completed. The experimental results showed that the feature points of the extracted image are extracted after edge extraction, some pseudo-feature points are removed, fewer and more accurate feature points can be obtained, and a large amount of noise can be suppressed to obtain more accurate information needed for matching. In order to evaluate the accuracy of the three-dimensional scene model reconstructed by this method, this paper evaluates the quality of three-dimensional model reconstruction of green plant landscape in three-dimensional scale with reference to the precise three-dimensional reconstruction model based on laser scanning point cloud. The results showed that the three-dimensional reconstruction model constructed by this method has high precision. The test of Hausdorff distance of green plants is between 0 and 3mm, which can meet the requirements of three-dimensional model reconstruction of green plant landscape on both sides of highway.

9 REFERENCES

►Ohtake, T., Kimijima, I., Fukushima, T., et al. (2015). Computer-assisted complete three-dimensional reconstruction of the mammary ductal/lobular systems: implications of ductal...


