Research on Image Visual Information Coordinate Transformation Model Algorithm

Liping Yuan
School of Information Engineering, Wuhan HuaXia University of Technology, Wuhan 430223, China.
Email: 94524948@qq.com

Abstract. 80% of the information obtained by human beings comes from vision, using pictures and videos as information carriers, and the information obtained from them is called visual intelligence. Visual intelligence analysis refers to extracting the size, distance, speed and other parameter information of the target object from the picture, video. In real life, some measuring tools can be used to get the size, distance and speed of the object Equal parameter information, but only through the picture is not directly available Arbitrary The actual distance between the two points needs to establish a transformation model between the plane coordinate system of the image and the actual coordinate system of the image. The mapping between the actual coordinate system and the phase plane coordinate system is established, the mapping relationship between the pixel plane and the pavement coordinate plane and the imaging model of the monocular camera are established. The vanishing point algorithm is used to calculate the distance from the focus to the center of light and the angle between the optical axis and the ground To achieve Based on the picture Picture actual distance between the target objects to facilitate the intelligence analysis of various picture information.

1. Introduction
Human beings obtain a variety of information to distinguish and treat different things through their own senses as information is ubiquitous, and visual is the main information acquisition methods. The information intelligence obtained from pictures and videos is defined visual intelligence, which refers to information such as the size, distance, and speed of target objects extracted from pictures and videos. Visual intelligence analysis plays a particularly important role in various popular areas such as drone reconnaissance and computer vision[1]. Although researchers have studied some visual systems to obtain relevant information, however it can only be analyzed by a picture or a video under lots of circumstance. The actual distance between the points cannot obtain only through the picture. A transformation model needs to be established between the plane coordinate system of the image and the actual coordinate system, a mapping relationship between the actual coordinate system and the phase plane coordinate system , So that each point in the actual coordinate system can have a corresponding coordinate in the phase plane coordinate system, this conversion relationship is the key to directly obtain the actual distance between the points to be measured through the picture.

2. Establishment Of Coordinate Transformation Model Based On Phase Plane and Actual Pavement

2.1. Camera Imaging Principle And Ranging Method
The image imaging is related to the parameters inside the camera according to the principle and properties of lens imaging. Different camera parameters determine the optical properties of the picture.
Three-dimensional spatial information can be transformed into pictures in two-dimensional space on the basis of the basic principle of a camera which is shown in Figure 1.

\[ f = \frac{y_0 - y}{D} \]  

Then the distance from the point \( P \) to the focus of the camera is

\[ D = \frac{hf}{y_0 - y} \]  

Figure 1 shows the camera's photographing position under ideal conditions. However, the camera usually has an inclined angle in real life. Formula (2) is used for distance calculation in an ideal state, and it can only calculate the distance from a single point to the camera, and cannot meet the distance measurement under actual conditions. In order to more accurately measure the distance between objects in the picture, we propose a ranging algorithm based on the transformation model of the phase plane and the actual road surface coordinate system.

2.2. Ranging Algorithm Based On Phase Plane And Actual Pavement Coordinate System Conversion Model

The distance measuring method [2] of the camera introduced in 2.1 can quickly find the coordinates of each pixel. However, the points in space are arbitrary in practice, and their corresponding relationship isn’t a linear relationship. Therefore, it is necessary to establish a transformation model between the phase plane coordinate system formed by the camera and the actual plane coordinate system which can give the position of certain point in phase plane, so the position of this point in actual coordinate system could be figured out. According to the phase plane coordinates of any two points in the picture, the coordinates in the actual coordinate system can be obtained, and then the actual distance of the object can be obtained.

The camera lens is regarded as a pinhole imaging [3], and the transformation of the phase plane coordinate system to the road surface coordinate system requires the transformation of the tilt coordinate system and the camera coordinate system. The mapping relationship is shown in Figure 2.
Figure 2. Mapping relationship between pavement coordinate system and map plane.

$X_c Y_c Z_c$ is the camera coordinate system, $X_{cs} Y_{cs} Z_{cs}$ is the rotation coordinate system formed by the rotation around the axis $Z_c$, $X_w Y_w Z_w$ is the actual road surface coordinate system, and $uv$ is the image plane coordinate system. The four coordinate systems are related to each other. After corresponding mathematical transformation [4], the mapping relationship between the road surface coordinate system and the phase plane coordinate system can be obtained.

The mathematical relationship between the simplified coordinate system $uv$ and the road surface coordinate system $X_w Y_w Z_w$ is equation (3),

$$
\begin{bmatrix}
    u \\
    v \\
    1
\end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix}
    k & 0 & R & T \\
    0 & 1 & 0 & 0
\end{bmatrix} \begin{bmatrix}
    X_w \\
    Y_w \\
    Z_w \\
    1
\end{bmatrix}
$$

(3)

The matrix $k$ in equation (3) is called the internal matrix, and its value is only related to the internal parameters of the camera; the matrix $R$ is called the rotation matrix, and its value is related to the rotation angle $\psi, \phi, \theta$; the matrix $T$ is called the translation matrix, which is the phase plane translation to the road surface coordinate system distance. According to equation (3), $X_w, Y_w, Z_w$ and the coordinate transformation with any point of the phase plane can be obtain through inverse transformation. Since the road surface is a plane coordinate and has nothing to do with height, let $Z_w = 0$, we have

$$
\begin{align*}
X_w &= \frac{l[u_s \cos(p) \sin(t) + v_s \sin(p)]}{v_s \cos(t) + f_s \sin(t)} \\
Y_w &= \frac{l[-u_s \sin(p) \sin(t) + v_s \cos(p)]}{v_s \cos(t) + f_s \sin(t)}
\end{align*}
$$

(4)

In Equation (4), the expressions of $u_s$ and $v_s$ are
\[
\begin{align*}
    u_s &= (u-u_0) \cos(s) + (v-v_0) \sin(s) \\
    v_s &= (u-u_0) \sin(s) - (v-v_0) \cos(s)
\end{align*}
\] (5)

Assuming that the positions of the two feature points in the phase-plane coordinate system are \(G(u_1,v_1)\), \(H(u_2,v_2)\) respectively, the coordinates of the two feature points in the actual road surface coordinate system can be calculated as \((X_{w1},Y_{w1}),(X_{w2},Y_{w2})\) by formulas (3) to (5), then the Actual distance is

\[
|Ax| = \sqrt{(X_{w1} - X_{w2})^2 + (Y_{w1} - Y_{w2})^2}
\] (6)

The transformation model of phase plane to actual road surface coordinates has been established, but three unknown quantities still need to be solved. Assuming the camera's placement position and various parameters are known, the value \(s\), \(p\), \(t\), \(l\) of the rotation angle can be obtained, and three unknown quantities can be solved. Then find the unknown parameters \(k\), \(R\), \(T\) by means of the vanishing point method when we can only get the picture data.

2.3. Value \(k\), \(R\), \(T\) Determination Based On Vanishing Point Method

According to the pinhole imaging principle, the camera projects a spatial object onto the projection surface to obtain a single-sided projection image that is closer to the visual effect. Objects close to the camera's origin are shown larger in the picture, and objects farther away are displayed smaller in the picture in the imaging plane, the two actual parallel lines that are not parallel will converge at a point, called the vanishing point [6]. This point is also the projection of an infinite point on the imaging plane, as shown in Figure 3.

![Figure 3](image)

**Figure 3.** The vanishing point of the actual road straight mapping to the camera coordinate.

In Figure 3, the left side is the two parallel lines \(AB\parallel CD\), \(AC\parallel BD\) in the road surface coordinate system. After shooting by the camera, the straight line \(AB\) in the camera coordinate system is no longer parallel to \(CD\), and it will intersect at the point \(L\) after extension. Meantime line \(BC\) and \(AD\) will intersect at point after extension, these two points of intersection are vanishing points.

2.3.1. Determination of the rotation matrix \(R\). The three orthogonal lines in the space coordinate system are also orthogonal in the imaging. For example, the actual road surface coordinate system is mapped to the three straight lines \(l\), \(m\), \(n\) in the camera coordinate system.

\[
\vec{l} \times \vec{m} = \vec{n}
\] (7)
\[
\begin{align*}
\langle l \times m \rangle &= 0 \\
\langle l \times n \rangle &= 0 \\
\langle m \times n \rangle &= 0
\end{align*}
\]

(8)

Formulas (7) and (8) show the positional relationship of the three straight lines in the camera coordinate system. Let the vanishing points correspond to the three straight lines \( L, M, N \). In the camera coordinate system, their coordinates are \( L(X_{cl}, Y_{cl}, -f), M(X_{cm}, Y_{cm}, -f), N(X_{cn}, Y_{cn}, -f) \), the relationship between the three points can be expressed by the following equations:

\[
\begin{align*}
X_{cl}X_{cm} + Y_{cl}Y_{cm} + f^2 &= 0 \\
X_{cl}X_{cn} + Y_{cl}Y_{cn} + f^2 &= 0 \\
X_{cm}X_{cn} + Y_{cm}Y_{cn} + f^2 &= 0
\end{align*}
\]

(9)

During the imaging process, the coordinate transformation of the projection point is only related to the placement of the camera. Let \( s \) as the coordinates of the vanishing point be in the road coordinate system, \( s' \) is the coordinates of the vanishing point in the camera coordinate system, and 

\[
R = \begin{bmatrix}
\frac{X_{cm}}{\sqrt{X_{cm}^2 + Y_{cm}^2 + f^2}} & \frac{X_{cl}}{\sqrt{X_{cl}^2 + Y_{cl}^2 + f^2}} & \frac{X_{cn}}{\sqrt{X_{cn}^2 + Y_{cn}^2 + f^2}} \\
\frac{Y_{cm}}{\sqrt{X_{cm}^2 + Y_{cm}^2 + f^2}} & \frac{Y_{cl}}{\sqrt{X_{cl}^2 + Y_{cl}^2 + f^2}} & \frac{Y_{cn}}{\sqrt{X_{cn}^2 + Y_{cn}^2 + f^2}} \\
-\frac{f}{\sqrt{X_{cm}^2 + Y_{cm}^2 + f^2}} & -\frac{f}{\sqrt{X_{cl}^2 + Y_{cl}^2 + f^2}} & -\frac{f}{\sqrt{X_{cn}^2 + Y_{cn}^2 + f^2}}
\end{bmatrix}
\]

(10)

The rotation matrix \( R \) can be obtained from equation (10).

2.3.2. Determination of the internal matrix \( R \). Assuming that the three-point coordinates of \( L, M, \) and \( N \) are known and the inverse solution is obtained from equation (9), the effective focal length \( f \) can be obtained.

\[
f = \sqrt{- (X_{cl}X_{cm} + Y_{cl}Y_{cm})} = \sqrt{- (X_{cl}X_{cn} + Y_{cl}Y_{cn})} = \sqrt{- (X_{cm}X_{cn} + Y_{cm}Y_{cn})}
\]

(11)

According to equation (11), it can be known that the coordinates of two points in a known camera coordinate system can calculate the effective focal length \( f \) of the camera, and the value of the internal matrix \( k \) can be obtained by being close.

2.3.3. Determination of translation matrix \( T \). As shown in Figure 4, the position of the camera's
optical center is \( o \), \( AB \) is an actual straight line on the road surface. \( A_iB_i \) is a imaging on the image plane of straight line \( AB \). \( v \) is a vanishing point of \( A_iB_i \) and another straight line, we draw a line \( A_iB_2 \) parallel to \( AB \) on \( A_i \).

![Figure 4](image)

**Figure 4.** Schematic diagram of translation vector solution.

According to the optical knowledge and the similar triangle theorem \(^7\),

\[
\frac{|A_iB_2|}{|OV|} = \frac{|A_iB_1|}{|VB_1|} \Rightarrow |A_iB_2| = \frac{|A_iB_1||OV|}{|VB_1|}
\]

\[
\frac{|OA|}{|OA_2|} = \frac{|AB|}{|A_iB_2|} \Rightarrow |OA| = \frac{|AB||OA_2|}{|A_iB_2|} = \frac{|AB||OA_2|}{|A_iB_1||OV||VB_1|}
\]

(12)

From equation (12), it is known that \( A \) is the origin of the road surface coordinate system, and the vector \( OA \) is the required translation vector. In the process of solving the model, the value \( OA \) is solved by knowing the length \( AB \) of the reference object, that is,

\[
l = \frac{|AB||OA_2|}{|A_iB_1||OV||VB_1|}
\]

(13)

In Equation (13), \( l \) is the distance from the origin \( o \) of the phase plane to the origin of the actual pavement coordinate system, the matrix \( T \) can be derived by passing the points in the selected pavement coordinate system and the corresponding points in the image plane coordinate system.

2.3.4. Determination of the third vanishing point. In the real state, it is difficult to obtain the third vanishing point directly. The third vanishing point needs to be obtained through mathematical relationships. As shown in Figure 3, the four points in the set ABCD are the four points in the actual road surface, and \( AB // CD, AD // BC \), from the definition and properties of the vanishing point, two parallel straight lines are on the phase plane. In the formation of two vanishing points. In Figure 3, \( L \) is the vanishing point of the two straight lines \( AB \) and \( CD \), and \( M \) is the vanishing point of the two straight lines \( AD \) and \( BC \). The coordinate system is established by using the lower left corner of the phase plane as the origin. The coordinates of the four points of \( ABCD \) are \((x_L, y_L), (x_M, y_M), (x_C, y_C), (x_D, y_D)\). From the coordinates, the equations of the four straight lines in the phase plane coordinate system can be obtained. After calculation, we can know the equations of the four straight lines \( AB, CD, AD, BC \) in equations (14) to (17).

The linear equation is \( AB \):
\[ y_i - y_A = \frac{y_B - y_A}{x_B - x_A} (x_i - x_A) \] (14)

The linear equation is BC:
\[ y_2 - y_B = \frac{y_C - y_B}{x_C - x_B} (x_2 - x_B) \] (15)

The linear equation is CD:
\[ y_3 - y_C = \frac{y_D - y_C}{x_D - x_C} (x_3 - x_C) \] (16)

The linear equation is DA:
\[ y_4 - y_D = \frac{y_A - y_D}{x_A - x_D} (x_4 - x_D) \] (17)

Let the coordinates of the two points \( L, M \) be \((x_L, y_L), (x_M, y_M)\) respectively, as can be seen from 2.3.2, \( u_0, v_0 \) represent the optical center, which can be regarded as the origin of the camera coordinate system, then the new coordinates of the two points \( L, M \) in the camera coordinate system will be \((u_0 - X_L, v_0 - Y_L, -f), (u_0 - X_M, v_0 - Y_M, -f)\) respectively. Let the third vanishing point be \( N \), the coordinates in the camera coordinate system be \( N(u_0 - X_L, v_0 - Y_N, -f) \), and substitute the coordinate values of the two points \( L, M \) in the camera coordinate system into formula (9), then the system of equations is
\[
\begin{cases}
(u_0 - X_L)(u_0 - X_N) + (v_0 - Y_L)(v_0 - Y_N) + f^2 = 0 \\
(u_0 - X_M)(u_0 - X_N) + (v_0 - Y_M)(v_0 - Y_N) + f^2 = 0
\end{cases}
\] (18)

From the equation (18), the coordinates of the point \( N \) in the phase plane coordinate system and the coordinates in the camera coordinate system can be obtained, thereby deriving the values of the internal matrix \( k \), rotation matrix \( R \), and translation matrix \( T \). Calibration and ranging issues. Therefore, the algorithm flow based on the vanishing point method to solve the transformation model of picture visual information coordinate system is shown in Figure 5.
3. Solution And Analysis Of Results Based On Phase Plane And Actual Pavement Coordinate Conversion Model

Through data searching [8], given the static image shown in Figure 6, the distance between the two pillars of the Beijing-style guardrail in the picture is 3.080m, and the parameter of the red vehicle A is $4342 \times 1840 \times 1500$mm. Due to various problems such as shooting equipment and the environment, the quality of the imaging picture is poor, which seriously affects the accuracy of the detection of the target object. It is necessary to first filter the image and filter irrelevant noise to improve the quality of the picture.

![Figure 6. Static original image.](image1)

![Figure 7. Gaussian filtered image.](image2)

3.1 Image Preprocessing

In order to reduce the experimental error, Gaussian filtering is performed on the picture to remove the noise of the picture, which facilitates the selection of the feature points of the picture. The filtering result is shown in Figure 7.
3.2. Selection Of Feature Points In Image Plane And Solution Of Vanishing Points

It can be known from the properties of the vanishing point that only the coordinates of the two vanishing points in the vertical and parallel directions need to be solved and the actual distance is measured to calculate the corresponding parameters.

The horizontal line and vertical line of the stone brick in the filtered picture shown in Figure 7 are perpendicular to each other, and the calculation mode of the vanishing point is satisfied. Select six points A, B, C, D, E, and F in Figure 7 to calculate the camera parameters, as shown in Figure 8. The actual length of EF, the guardrail column is 3.080m. Since AB // CD, AD // BC in the pavement coordinate system, the selection of A, B, C, and D meets the conditions for the vanishing point. EF is the selected known actual distance.

![Figure 8. Selection of feature points.](image)

Using Photoshop software to calibrate the feature points, the coordinates of each point are obtained as shown in Figure 8. The origin of the coordinates is F point, then
According to the A, B, C, and D coordinate values, the analytical equations of the four straight lines $AB$, $BC$, $CD$, $DA$ in the image plane can be obtained as

\[
\begin{align*}
AB: y &= 0.2150x + 686.7250 \\
BC: y &= 0.8519x - 26.5000 \\
CD: y &= 0.1127x + 891.1892 \\
AD: y &= -0.9915x + 12038
\end{align*}
\] (19)

From the analytical formula of equation (19), we can figure out the intersection point $M(X_m, Y_m)$ of $AD$ and $BC$ as well as the intersection point $N(X_n, Y_n)$ of $AB$ and $CD$, the coordinates of two vanishing points are determined, and then the vanishing point algorithm is used to find the third The vanishing point $L(X_l, Y_l)$ values, that is, the coordinates of the three vanishing points are as follows:

\[
\begin{align*}
M(X_m, Y_m) &= (1292.32, 1047.5) \\
N(X_n, Y_n) &= (-1591.2, 704.41) \\
L(X_l, Y_l) &= (-345.09, 2690.5)
\end{align*}
\] (20)

The basic pixel values $p_x$, $p_y$, $u_0$, $v_0$ from the pixel data in Figure 8 are shown in Table 1.

| Project Name | $p_x$ | $p_y$ | $u_0$ | $v_0$ |
|--------------|------|------|------|------|
| Value / Pixel | 2000 | 1500 | 1000 | 750 |

The coordinates of the vanishing point in the camera coordinate system are shown in Table 2.

| Project | $x_c$ | $y_c$ | $z_c$ |
|---------|------|------|------|
| First correspondence Point | 292.32 | 297.5 | -878.69 |
| Second correspondence Point | -2591.2 | -45.59 | -878.69 |
| Third correspondence Point | -1345.09 | 2540.5 | -878.69 |

The value of the effective focal length $f$ can be obtained by using formula (11), that is,

\[
f = \sqrt{-(X_c \cdot X_{cm} + Y_c \cdot Y_{cm})} = 878.69 \text{ pixel}
\]

Using the coordinates of the vanishing point in the camera and the inverse transformation formula (12), and Using the principle of solving the translation vector, as shown in Equation (13), the distance from the origin of the phase plane to the origin of the road surface is

\[
l = \frac{|AB| |OA|}{|A_i B_i | |OV|} = 5.16 m
\]

According to the rotation matrix, the values of the transformation angles of the three coordinate systems can be derived, as shown in Table 3.
Table 3. Conversion angle.

| Name of the angle | Internal declination $s$ | Declination $p$ | Elevation angle $t$ |
|------------------|--------------------------|----------------|------------------|
| Value (Radian)   | 0.1910                   | 1.6463         | 0.2349           |

According to equations (4) ~ (6), we can know that the actual distance between the two feature points is

$$|\Delta x| = \sqrt{(X_{w1} - X_{w2})^2 + (Y_{w1} - Y_{w2})^2} = 17.36m$$

3.3. Determine The Actual Distance Of Picture Information Reference Points

It is assumed that the position of the photographer at point B is shown in Figure 9, and the coordinates of A, B, C, and D are marked in Figure 9.

![Figure 9. Reference point pixel position.](image)

Then the analytical equations of the straight line $AB$ and the straight line $CD$ are:

\[
\begin{align*}
AB & : y = 0.2242x + 1062.20 \\
CD & : y = -0.1034x + 729.24
\end{align*}
\]

The coordinates of the intersection $Q$ of the two lines are $(-1002.24, 832.91)$ According to the transformation relationship of the coordinate system, we can know the coordinates of the intersection $Q$ and the point $B$ in the road surface coordinate system is $Q(18.12, -7.02), B(6.45, -7.82)$.

In the actual situation, the straight line $AB$ is perpendicular to the straight line $CD$, so the distance between the intersection and the point is the distance of the photographer from the left border of the road:

$$|\Delta x| = \sqrt{(X_{wQ} - X_{wB})^2 + (Y_{wQ} - Y_{wB})^2} = 11.70m$$

That is, the actual distance between the target objects in the picture is calculated based on the picture. The algorithm is useful for extracting various parameter information such as the size, distance, and speed of the target object for various pictures, and then performing intelligence analysis.

4. Conclusion

With the advent of the era of big data, information leakage incidents have emerged endlessly. Not only must we improve our access to effective information, but we must also learn to conceal effective information accordingly to protect ourselves, the motherland, and the homeland. The biggest benefit is visual information information leakage There are a lot of proportions. The article focuses on the
acquisition of picture information. According to the conversion model between the plane coordinate system of the picture and the actual actual coordinate system, the vanishing point algorithm is used to measure the actual distance between the target objects in the picture. Conducive to the intelligence analysis of various picture information.

5. Reference

[1] Gong Ke, Secrets of photo leaks [J]. 2016, Oil Knowledge, pp. 35.
[2] Lai Jiawei, He Yuqing, Li Xiaopeng, et al. Design of manipulator target positioning system based on monocular vision [J]. 2019, Optical Technology, pp. 6-11.
[3] Sui Mingming, spatial information extraction and error analysis of single close-range images [J]. 2013, Nanjing Normal University, pp. 152.
[4] Deng Sibin, Research on Traffic Video Anti-shake Algorithm for Complex Motion Scenes [D]. 2017, University of Science and Technology of China, pp. 103.
[5] Liu Xuejun, Wang Meizhen, Yan Yan et al. Research Progress on Geometry of Single Image [J]. 2011, Journal of Wuhan University (Information Science Edition), pp. 941-947.
[6] Guo Buyu, Yu Jia, Wang Jiao et al. Binocular vision for fish fry size measurement [J]. 2017, Optics Technology, pp. 153-157.
[7] Luo Guie, Binocular stereo vision depth perception and several problems of 3D reconstruction [D]. 2012, Central South University, pp. 131.
[8] Kong Fanqi, A study of UAV range measurement and obstacle avoidance methods based on binocular vision [J]. 2019, Science and Technology Wind, pp. 101.