A Study on Positioning Algorithm for the Abnormal Temperature Area of Cotton Warehouses Based on Infrared Temperature Measurement and Visible Light Fusion

Xinran Zhao¹,², Yao Pan³ and Xuxu Zhang²

¹China University of Mining and Technology (Beijing) 100083
²China Academy of Safety Science and Technology, Beijing 100012, China
³Northeast normal university, Changchun 100024, China

Abstract. Aiming at the positioning problem in the fire emergency response of cotton warehouses, a positioning algorithm for the abnormal temperature area in cotton warehouses is proposed in this paper based on infrared temperature measurement and visible light fusion. This study combines image data fusion, temperature, and gas sensor data to detect the abnormal temperature area, and positions the spatial three-dimensional point through camera calibration and pose estimation, to achieve early warning and emergency response of cotton warehouses. The experiments show that this algorithm can effectively achieve the positioning of the abnormal temperature area.

1. Introduction

China is a country with a large cotton reserve, and our maximum stock in 2016 was 12 million tons. Due to flammable, smoldering and spontaneous combustion properties of cotton, fire accidents occur frequently in the process of cotton storage, which have caused a huge loss. The possible causes of cotton fires mainly include: lightning strike, static electricity, smoldering, and external fire sources. The sources of danger causing fires in cotton warehouses mostly occur in the transportation or storage operations of cotton. The time from a spark entering some cotton packet, to the occurrence of smoldering is quite long, and the cotton is in the stacked and closed status after put into a storage warehouse, so it is difficult to find the fire point, and even more difficult to judge the status before fire. Therefore, it is very necessary to find out the position of the abnormal temperature point through technical means and realize early warning and emergency response to avoid a huge economic loss.

Because of the complex situation of cotton warehouses, multiple sensors are required to obtain multi-directional features of the target simultaneously to achieve fire warning. However, judging only according to sensor signals has the disadvantages of low accuracy and high false warning rate. In this case, the target can be accurately detected only through the effective fusion analysis of sensor signals.

A decision tree is a common multi-source information coupling algorithm, which can bring great convenience to the establishment of a system. After the information is coupled, it is necessary to find out the abnormal temperature point. The premise of realizing the positioning algorithm is to accurately calibrate the camera. No matter it is in image measurement or in the machine vision application, the calibration of camera parameters is a very critical link. The accuracy of the calibration results and the stability of the algorithm directly affect the accuracy of the results generated by the camera work. At present, there are many mature theories regarding camera calibration, such as the direct linear
calibration method, nonlinear camera calibration method, self-calibration method, etc. [1]. However, in practical engineering, few calibration tools can be directly used for camera calibration quickly and easily. The Zhang’s Calibration Method was adopted in this study, which is a practical method of using a plane checkerboard for camera calibration proposed by Dr. Zhengyou Zhang in a paper published on the ICCV, an international top conference in 1999 [2]. This method combines the photography calibration method and the self-calibration method, overcoming the shortcoming of the former that requires the target to be a high-accuracy three-dimensional calibration object and solving the problem of poor robustness of the latter. It can realize calibration by only using a printed checkerboard and taking a few sets of pictures from different directions. The user can make the calibration pattern himself / herself, which is not only practical and flexible, but also highly accurate and robust. Therefore, the method was quickly adopted by the whole world, greatly promoting the progress of 3D computer vision from the laboratory to the real world.

After the accurate calibration of the camera was realized, the pose estimation was used to realize the three-dimensional positioning algorithm. In the pose measurement process, only one camera was used, which has a large measurement range and a simple structure, and widely used in robot tracking and navigation, aircraft docking, and aircraft attitude control [3]. Monocular Vision Positioning is to use the feature point on the positioning target to estimate the pose. Based on the mathematical measurement model of the camera, the topological projection relationship between the spatial target’s feature point and its corresponding image coordinate point can be established. Then, how to further solve the relative pose between the spatial target and the measuring camera is the key in the monocular vision pose measurement, and is also the focus of the research of visual positioning technology. The calculation models in pose measurement are generally non-linear, so a mathematical manipulation should be performed according to the features of the pose measurement model, in order to achieve a linear solution for pose positioning. The Lie algebraic manipulation group was locally linearized using the exponential mapping of the Lie group to establish a linear mapping relationship between the image displacement of the feature point and the spatial three-dimensional pose manipulation to achieve the linear solution of pose measurement [4-5]. In addition, the algebraic manipulation according to the features of the pose measurement model can also achieve the linear solution of the pose. Wu Fuchao [6] et al. analyzed the problem of the feature point by using the two images under translational motion, and proved that four points can achieve linear solution of pose measurement. In the literature [7], the three-point pose measurement model was transformed into an eighth-order equation containing only even-order terms and constant terms. In the process of pose solution for four or more points, three even-order terms are used as variables to realize the linearization of the pose measurement model. In literature [8], the linear solution of the pose measurement of five feature points was realized. In this study, we use the coplanar P4P in PNP algorithm to estimate the pose, and the pose of the camera is obtained and the spatial position of the abnormal temperature point is calculated.

The main content of this paper is to judge whether there is an abnormal temperature point by coupling the existing image containing flame position information and the data obtained from multiple sensors. If the abnormal temperature point is detected, its spatial coordinates should be positioned. Firstly, the multi-source information is coupled through the decision tree; secondly, the camera is calibrated with the Zhang’s Calibration Method, and finally the camera pose is estimated through the P4P algorithm, and the spatial pose of the abnormal temperature point is obtained.

2. Principles
To determine whether the image containing the position information of an abnormal temperature area is real, it should be analyzed by conducting multi-source coupling of the information. In this paper, the data from multiple sensors involving temperature, oxygen, carbon monoxide and carbon dioxide is used to judge whether there is an abnormal temperature area. If the abnormal temperature area exists, its coordinate position of the temperature abnormal area in space can be found through the output image. The overall algorithm flow chart is as follows:
2.1. Multi-source Information Coupling

In this study, the decision tree model based on information entropy is used for the information coupling. The decision tree algorithm is a non-parametric decision-making algorithm. It performs multi-level classification and judgment according to different features of data, and finally makes the required prediction results. This algorithm can solve both the classification algorithm problem and the regression problem, with a good explanatory ability. For different intentions there are multiple methods of constructing a decision tree. The intention to construct a decision tree mainly depends on the classification of dimensions of each decision point of the decision tree and the threshold nodes of these dimensions and other details.

Information Entropy is the most common method in the construction of a decision tree. The meaning of entropy in the information theory is the uncertainty of random variables. The greater the entropy is, the greater the uncertainty will be; similarly, the smaller the entropy is, the smaller the uncertainty will be. The information entropy is a measure of the amount of information required to eliminate the uncertainty, namely, the amount of information that an unknown event may contain. The general principle of dividing the information entropy is to make the entropy value of the overall decision-making system smaller and smaller, so that the system develops in a more and more certain direction. The formula for information entropy is:

$$H(X) = - \sum_{x \in X} P(x) \log P(x)$$

Where, x is any random variable and H(X) is its information entropy.

This paper couples the image with the data from multiple sensors to judge whether there is an abnormal temperature point according to the sorting of number of categories after classification.

2.2. Calibration of Visible Light and Infrared Cameras

To restore the position of the object imaged by the camera in the real world, it is necessary to know how the object in the real world changes to the computer image plane. Therefore, camera calibration is required. The camera calibration is the process of solving camera parameters. One of the purposes of camera calibration is to make clear the transformation relationship between the world object and the image plane, and to solve the internal and external parameter matrices. Secondly, the perspective projection of the camera can produce distortion. Another purpose of the camera calibration is to solve the distortion coefficient and then used for image correction. The flow chart of camera calibration is as follows.
2.2.1. Collection of Calibration Sample Pictures.

2.2.2. Principles of Camera Calibration. The camera calibration involves four coordinate systems, namely the world coordinate system, the camera coordinate system, the image coordinate system and the pixel coordinate system. The world coordinate system is a three-dimensional coordinate system used to describe the position of the target in the real world; the camera coordinate system is a coordinate system established on the camera, which can describe the position of the object relative to the camera; the image coordinate system is used to describe the projection and transmission relationship of the object from the camera coordinate system to the image coordinate system in the imaging process; and the pixel coordinate system is used to describe the coordinate of the image point of the object on the digital image after the object is imaged, and it is the coordinate system where the information actually read from the camera is located.

The transformation from the world coordinate system to the camera coordinate system can be achieved by rotation and translation. The relationship between the two is as follows:

\[
\begin{bmatrix}
X_c \\
Y_c \\
Z_c \\
1
\end{bmatrix}
= R \begin{bmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{bmatrix}
\]

Where, \((X_c, Y_c, Z_c)\) and \((X_w, Y_w, Z_w)\) denote the homogeneous coordinates of a point P in space in the world coordinate system and the camera coordinate system, respectively. R denotes the rotation matrix, and T denotes the translation matrix.

In the transformation from the camera coordinate system to the image coordinate system, the two has the perspective projection relationship, which can be expressed as:

\[
\begin{bmatrix}
X_c \\
Y_c \\
Z_c \\
1
\end{bmatrix}
= f \begin{bmatrix}
0 & 0 & 0 \\
0 & f & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{bmatrix}
\]

Where, \((x, y)\) is the homogeneous coordinates of a point P in space in the image coordinate system, and f is the focal length of the camera.

The transformation from the image coordinate system to the pixel coordinate system is the transformation from the image physical coordinate system to the pixel coordinate system. The transformation relationship is as follows:

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

Where, \((u, v)\) is the homogeneous coordinates of a point P in space in the pixel coordinate system, \((u_0, v_0)\) is the central pixel coordinates on the image, and the physical length of each pixel point in the image is \(d_x\) on the coordinate axis X and \(d_y\) on the coordinate axis Y, respectively.

In summary, all the transformation from the world coordinates to the pixel coordinates is:
Where, $\mathbf{M}_1$ is the internal parameter matrix of the camera, and it is determined by the internal parameters of the camera; $\mathbf{M}_2$ is the external parameter matrix of the camera, and it is determined by the orientation of the camera relative to the world coordinate system.

2.3. Three-dimensional Image Positioning

In this study, pose estimation was used for the three-dimensional positioning of the image based on the PNP (Perspective-n-Point) algorithm. The PNP algorithm is: to estimate the pose of the camera when the real coordinates of $N$ spatial points in the world coordinate system and the projection of these spatial points on the image are known.

Suppose that the camera is located at point A, when $N=1$, there is only one feature point $P_1$. Assume it is in the center of the image, then obviously the vector $\mathbf{AP}_1$ is the Z axis in the camera coordinate system, and the camera is always facing $P_1$, so the possible position of the camera is on the spherical surface with $P_1$ as the center of the sphere, but the radius of the sphere cannot be determined, so there are countless solutions.

When $N=2$, it is obvious that $\mathbf{AP}_1\mathbf{P}_2$ forms a triangle. Since the positions of $P_1$ and $P_2$ are determined, the side $P_1P_2$ of the triangle is determined, too, the direction angle of the vectors $\mathbf{AP}_1$ and $\mathbf{AP}_2$ from point A to the feature point can also be determined. Therefore, the length $r_1$ of $\mathbf{AP}_1$, the length $r_2$ of $\mathbf{AP}_2$ can be obtained by calculation. Thus, two spheres are obtained: sphere $\mathbf{B}$ with the center of $P_1$ and the radius of $r_1$; sphere $\mathbf{C}$ with the center of $P_2$ and the radius of $r_2$. The camera is located at the intersection of sphere $\mathbf{A}$ and sphere $\mathbf{B}$, so there are still countless solutions.

When $N=3$, there is an additional sphere $\mathbf{D}$ with $P_3$ as the center of sphere. The camera is located at the intersection of the three spheres, i.e. $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$. There will be 4 solutions, one of which is the true pose of the camera.

When $N>3$, the positive solution can be obtained. In order to solve the problem more quickly, while saving computer resources, four sets of solutions are obtained by using 3 points, and four rotation matrices and translation matrices are obtained. Then the world coordinates of the fourth point are substituted to obtain four projections in the image, of which, the one with the smallest projection error is namely the required positive solution.

In this paper, the P4P algorithm is used to solve the problem. After the camera pose is obtained, the coordinates of their intersection point can be solved according the principle “a point can be determined by two straight intersecting lines” as long as the equation of the two intersecting straight lines is known in the two-dimensional plane, as shown below:

![Figure 3. Two-dimensional Positioning Diagram the Camera.](image-url)
A and B during two shots can be estimated, and the coordinates of the camera will be also known. Then Points A and P, as well as points B and P are connected to form straight lines respectively, then, two straight line equations will be obtained. After an equation set is formed with the two equations their intersection points can be obtained through solution, the result is namely the coordinates of point P.

In three-dimensional space, the principle is the same as that in the two-dimensional space, except that two straight lines in two-dimensional space become two spaces in three-dimensional space. By solving P4P, the spatial positions A and B of the camera's two shots are obtained. Based on the coordinates of point P in the image, we can know which direction the point P is relative to the camera in space, which means we can determine a ray from the camera to point P. Two images are used to determine two rays about P, then the equation was solved to obtain the coordinates of the intersection point, thus, the spatial coordinates of P are obtained.

The information of two positions is coupled: (the precision is high, average may be taken).

3. Experiment and analysis

In the experiment, Intel (R) Core (TM) i7-855OU CPU @1.80GZ 2.00GHZ processor, 64-bit operating system, python3.6, VS2017 and OpenCV3.4 were used to implement the algorithm. In this paper, the sensor and image data obtained by experiments are coupled with python language. The image containing the flame position information that has been obtained is shown in Fig. 4 and the pixel coordinates of the flame area are output.

![Figure 4. Image with flame position information.](image)

Through the Zhang's Calibration Method, 20 images of checkerboard calibration board with different angles are required to calibrate the camera. The images of the checkerboard calibration board for infrared and visible light are shown in Fig. 5.

![Figure 5. Images of the checkerboard calibration board for visible and infrared lights.](image)

After the camera parameters are obtained, the cameras in two directions are required for pose estimation, and the images taken by the cameras in different directions are shown in Fig. 6.
Figure 6. Images taken by the cameras in two directions.

Figure 7 is a design idea diagram, in which, the directions of the X and Y axes are shown; points 1, 2, 3, and 4 denote P1, P2, P3 and P4, with point P1 as the origin of the world coordinate system; the straight line from 1 to 3 is X axis, the straight line from 1 to 2 is Y axis, and points A and B are the points to be measured. The pixel coordinates and the world coordinates of points P1, P2, P3 and P4 in the two images are known, and the pixel coordinates of points A and B are known as well, so the world coordinates of points A and B can be obtained.

Figure 7. Design Idea Diagram.

This paper only takes the coordinates of the X and Y directions in the world coordinate system. The world coordinates of point A are (-51, 18) (in cm), the coordinates determined by using the algorithm are (-50.449, 19.392), and the world coordinates of point B are (-55, 39), the coordinates measured by the algorithm are (-55.662, 40.624). The algorithm has a high accuracy and can be applied in actual scenes.

4. Conclusion
In order to realize emergency response in case of fire in a cotton warehouse, this paper proposes an emergency response technology for the abnormal temperature area of the cotton warehouse based on the fusion of infrared temperature measurement and visible light. The algorithm first couples the image and the data from multiple sensors involving temperature, oxygen, carbon monoxide, and carbon dioxide based on the decision tree to determine whether there is an abnormal temperature point; then, positions the abnormal temperature point through the image, during which, the camera is calibrated first to obtain internal and external parameters as well as distortion, and then the pose of the camera is obtained through the P4P algorithm to determine the spatial coordinates of the point to be solved. The algorithm has a high accuracy and can be applied in actual scenes.
Acknowledgments
This work was supported by National Key Research and Development Program of China (2017YFC0805900), China academy of safety production science basic business expenses special funds for scientific research projects (2015BAK16B-jy003), (2016JBKY13) (CXYF201902).

References
[1] De Ma S. A self-calibration technique for active vision systems [J]. IEEE Transactions on Robotics and Automation, 1996, 12 (1): 114-120.
[2] Zhang Z. Flexible camera calibration by viewing a plane from unknown orientations [C] // Proceedings of the seventh IEEE international conference on computer vision. IEEE, 1999, 1: 666-673.
[3] Dan Yu, Wei Wei, Yuanhui Zhang. Mobile robot following based on monocular vision [J]. Chinese Journal of Scientific Instrument, 2010 (3): 659-664.
[4] Drummond T, Cipolla R. Application of lie algebras to visual servoing [J]. International Journal of Computer Vision, 2000, 37 (1): 21-41.
[5] Ortegón-Aguilar J, Bayro-Corrochano E. Lie algebra and system identification techniques for 3D rigid motion estimation and monocular tracking [J]. Journal of mathematical Imaging and Vision, 2006, 25 (2): 173-185.
[6] Fuchao Wu, Zhanyi Hu. Linear algorithm for solving PnP problems [J]. Journal of Software, 2003, 14 (3): 682-688.
[7] Quan L, Lan Z. Linear n-point camera pose determination [J]. IEEE Transactions on pattern analysis and machine intelligence, 1999, 21 (8): 774-780.
[8] Tang J, Chen W S, Wang J. A novel linear algorithm for P5P problem [J]. Applied Mathematics and Computation, 2008, 205 (2): 628-634.
[9] Guangsheng Hu, Qing Wang, and Qingqin Shan. A study on the calibration method of infrared and visible dual cameras [J]. Technology Innovation and Application, 2017 (2017 04): 9-11.
[10] Guiwen Ren. Research on the calibration method of infrared and visible dual cameras based on OpenCV [J]. Science Technology and Engineering, 2016, 16 (3): 211-214.
[11] Jiawei Xing, Haifeng Tian and Fang Wang. Research on object pose estimation method of monocular camera [J]. Navigation, Positioning and Timing, 2019 (2019 04): 71-77.
[12] Yufei Zang, Yingzi Tan. Research on target pose estimation based on flexible control of rescue robot [J]. Industrial Control Computer, 2018 (6): 28.
[13] Fan Zhu, Fangsu Yu, Yiming Wu, et al. Precision analysis of P4P camera attitude calibration [J]. Acta Optica Sinica, 2018, 38 (11): 1115005.
[14] Peng Wang, Changku Sun, Zimiao Zhang. Linear solution for posture measurement in monocular vision [J]. Chinese Journal of Scientific Instrument, 2011, 32 (5): 1126-1131.