Research on the Fast Solution Method of Pilots' Head Attitude

Limin Zhang\textsuperscript{1}, Jie Jiang\textsuperscript{1,r,a}, Wei Fang\textsuperscript{1}, Kai Liu\textsuperscript{1}

\textsuperscript{1}Departments of Electronics and Science, University of Naval Aviation, Yantai 264000, China
\textsuperscript{r,hkd@tsg.edu.com}
Email: iamzlm@163.com, 81348541@qq.com, fangw4512@163.com, liukai4562@163.com

Abstract: Aiming at the problem of poor accuracy and poor real-time performance in the process of head pose calculation in pilot cabin simulation, a high accuracy and fast algorithm for head pose calculation is proposed. On the basis of traditional EPNP algorithm was improved, and combining with POSIT algorithm's speed advantage, select face five key feature points as virtual control points, according to the mapping relationship between space and virtual control point coordinates, establish a system of linear equations, then the internal camera parameters is calculated using Zhengyou Zhang camera calibration method, finally using POSIT algorithm of calculating to get the head rotation matrix and translation vector, by conversion from the head of Euler angle in real time. The experimental results show that compared with DLT, EPNP, POSIT and other traditional methods, the proposed method has higher accuracy, better real-time performance and better robustness in solving pilots' head attitude in continuous video.

1. Introduction
Cockpit simulated flight training is to provide pilots with real control elements, such as throttle, lever and pedal, and to restore real flight experience by combining the visual system and dynamic platform. It can conduct takeoff, landing, aerial refueling, formation flight, tactical confrontation and emergency response training in complex situations. Its low - base, low - risk, high - effect and high - return training mode is of great strategic significance to the future real training operations. By solving and tracking the pilot's head posture in cockpit simulation flight training, the functions such as head tracking aiming, attention allocation and fatigue detection can be realized, which is of great significance for improving the training level of pilots.

At present, the head attitude calculation method for pilots is still in the tradition of DLT\cite{1}, POSIT\cite{2} and EPNP\cite{3}.As a classical algorithm for perspective-N-Point (PNP) problems, DLT method directly constructs 12 unknown augmented matrices. On the premise of not considering the degrees of freedom of rotation matrices, it selects six pairs of control points, normalizes the mapping of each 3D control Point, and obtains an approximate rotation matrix by decomposition. POSIT algorithm, also known as proportional orthogonal projection iterative transformation algorithm, USES weak projection and linear system to estimate the rotation matrix and translation vector of the object on the premise of knowing 4 non-coplanar feature points. By repeatedly iterating the projection, the modified head attitude is obtained. EPNP algorithm is a non-iterative and closed PNP algorithm. On the premise of knowing 4 non-coplanar virtual control points and their 3D coordinates, the rotation vector is calculated by matching with corresponding points in the 2d camera coordinate system, and...
the rotation vector is converted into quaternions, and finally the euler Angle containing attitude information is calculated.

2. Head attitude theory and model
In the background of computer vision, the head attitude estimation is specified as the change of the head relative to the camera view direction, and the change of the head relative to the global coordinate system is inferred. It is necessary to know the parameters of the inherent camera in advance to eliminate the perception deviation caused by perspective distortion. The human head is generally regarded as a non-rigid object when solving the attitude calculation. Therefore, the head attitude can be represented by the pitch Angle, rolling Angle and yaw Angle through three degrees of freedom [4], as shown in Figure 1.

![Figure 1. Schematic diagram of head attitude Angle.](image)

The idea of head attitude estimation is to make the 2d projection of the 3D feature points on the model fit with the feature points on the plane image by rotating the 3D standard model, so as to obtain the rotation matrix and the translation vector [5], as shown in Figure 2.

![Figure 2. Image fitting effect picture.](image)

Generally, the face contains 68 key feature points, such as the tip of eyebrow, canthus, tip of nose, corner of mouth, and chin, etc. These key points are marked and their two-dimensional spatial coordinates are obtained [6]. Then the five feature points of left and right pupil, tip of nose and left and right corner of mouth were selected as the five virtual control points, as shown in Figure 3.

![Figure 3. Marking key feature points of face image.](image)

3. An improved EPNP - based approach to head attitude calculation
For the PNP algorithm, it aims to solve the camera pose from 3D-2D matching by determining the position and direction of the camera, giving the internal parameters of the camera and the n group
correspondence between 3D points and 2D projection [7]. It is mainly used in photogrammetry and computer vision, especially when dealing with camera tracking based on feature points, etc., hundreds of noise feature points usually need to be processed in real time, which requires efficient calculation methods [8].

To solve the face pose, EPNP algorithm needs to know at least four non-coplanar face feature points' THREE-DIMENSIONAL spatial coordinates. By matching the corresponding points in the two-dimensional camera coordinate system, the rotation vector can be calculated and converted into quaternions. Finally, the Euler Angle containing the attitude information can be calculated. The corresponding relationship between the virtual control points in the image plane and the world coordinate system is shown in Figure 4.

![Figure 4. Relationship between image coordinate points and virtual control points.](image)

The attitude settlement method based on the improved EPNP algorithm is based on the EPNP algorithm combined with the POSIT algorithm, which can not only ensure the accuracy of the solution, but also get the attitude data quickly. Firstly, five face feature points are selected as five non-coplanar virtual control points. According to the spatial relationship between the spatial coordinate points and the virtual control points, a linear equation group is established to calculate the virtual control point coordinates in the image coordinate system. Then, the internal parameters of the camera are calculated by Zhang Zhengyou camera calibration method [9]. Finally, the coordinates of the virtual control points are converted into rotation through the posit algorithm Transfer matrix and translation vector. The specific calculation process is as follows.

Firstly, the mathematical model of camera coordinate system [10] is established, and the formula is as follows:

\[
\begin{bmatrix}
X_i \\
Y_i \\
Z_i \\
1
\end{bmatrix} = \mathbf{K} \begin{bmatrix}
X_c \\
Y_c \\
Z_c \\
1
\end{bmatrix}
\]

(1)

Where \( \mathbf{K} = \begin{bmatrix} f & s & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \) is the internal parameter of the camera, \( \mathbf{R} = \begin{bmatrix} i^T \\ j^T \\ k^T \end{bmatrix} \) is the rotation matrix, \( t = (t_x, t_y, t_z)^T \) is the translation vector.

The relationship between image coordinate system \((x, y)\) and camera coordinate system \((X, Y, Z)\) is as follows:

\[
\begin{bmatrix}
x \\
y \\
1
\end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
\]

(2)
The relationship between spatial coordinate points and virtual control points in the world coordinate system is as follows:

\[ P^w_i = \sum_{j=1}^{5} a_j C^w_j, \quad i = 1, 2, ..., n \]  

(3)

Where \( \{C^w_j\} \) is the coordinate of the virtual control point in the camera coordinate system, and \([a_1, a_2, a_3, a_4, a_5]^T\) is the coordinate based on the control point in the space coordinate system.

With POSIT algorithm, rotation matrix \( R \) of the head is calculated according to the coordinates of the obtained virtual control points. Since the head attitude is three degrees of freedom, it is represented by a matrix of 3X3:

\[
R = \begin{pmatrix}
  r_1 & r_2 & r_3 \\
r_4 & r_5 & r_6 \\
r_7 & r_8 & r_9
\end{pmatrix}
\]

(4)

The rotation matrix \( R(\alpha) \), \( R(\beta) \), and \( R(\gamma) \) are calculated as follows:

\[
R(\alpha) = \begin{pmatrix}
  1 & 0 & 0 \\
  0 & \cos(\alpha) & \sin(\alpha) \\
  0 & -\sin(\alpha) & \cos(\alpha)
\end{pmatrix}
\]

(5)

\[
R(\beta) = \begin{pmatrix}
  \cos(\beta) & 1 & 0 \\
  0 & -\sin(\beta) & \cos(\beta) \\
  -\sin(\beta) & 0 & \cos(\beta)
\end{pmatrix}
\]

Finally, through the relationship between the rotation matrix and the rotation Angle, the head attitude Angle roll, pitch, yaw can be obtained:

\[
\text{roll} = \alpha = \arctan \left( \frac{-r_7}{r_5 \times c_2 + r_4 \times s_2} \right)
\]

\[
pitch = \beta = \arctan \left( \frac{r_5 \times s_2 + r_4 \times c_2}{r_4 \times c_2 - r_5 \times s_2} \right)
\]

\[
yaw = \gamma = \arctan \left( \frac{r_4}{r_1} \right)
\]

(6)

Through spatial coordinate mapping, EPNP has a high accuracy in solving virtual control points. In particular, 5 key non-coplanar feature points are selected to further improve the calculation accuracy. At the same time, combined with POSIT algorithm, only a few points of information are needed, which has the advantage of avoiding the process of solving complex nonlinear equations, greatly reducing the computational complexity and computing time, and at the same time, the robustness to errors and disturbances is enhanced.
4. Experimental verification

4.1. Calculate the camera's internal parameters
In the experiment, Logitech C920 camera with a resolution of 1920×1080 was used for image acquisition. Template images of different angles and sizes were taken, and the shooting results were shown in Figure 5.

![Figure 5. Template image shooting results.](image)

The internal parameters of the camera are solved, and the results are as follows:

Parameter matrix:

\[
\begin{bmatrix}
1.5581405e+03 & 0.00000000e+00 & 1.14318794e+03 \\
0.00000000e+00 & 1.57197780e+03 & 5.93763907e+02 \\
0.00000000e+00 & 0.00000000e+00 & 1.00000000e+00 \\
\end{bmatrix}
\]

dist : 

\[
\begin{bmatrix}
-0.3771424 & 2.4115141 & -0.0398826 & -0.0138733 & -4.757568 \\
\end{bmatrix}
\]

4.2. Compare the different solutions
DLT method, EPNP method, POSIT method and IEPNP method are compared from two aspects of calculation precision and calculation time respectively under different corresponding points. The range of corresponding points is [20,120] and the step length is 10. Error and time consumption of each independent point are obtained through 10,000 independent experiments, and the average value is finally taken. The experimental results are shown in Figure 6.
4.3. Actual scenario test

AFW data set [11] is selected as the test solution accuracy, which contains 205 images and 473 marked faces, each of which contains rectangular boundary boxes, 6 landmark data and attitude angles. The test results are shown in Figure 7.
In the cockpit simulation training, the real-time standard of solving the pilot's head attitude is no less than 30FPS. It is tested in an operating environment with hardware configuration of Intel Core i5-8250U CPU, MX150 graphics card and 12G memory. The average detection frame number can reach 66.56 FPS, which meets the real-time requirements. Compared with the traditional attitude solving method, the solving speed is significantly improved, and it is more robust to detection errors and external interference. The test results are shown in Figure 8:

5. Conclusion
In order to quickly solve the high-precision head attitude, based on the current EPNP algorithm with the highest resolution, and combined with POSIT algorithm, a new high-precision fast head attitude solving method is proposed for the pilot cockpit simulation training. The experimental results show that the algorithm has better precision and time than the traditional algorithm, meets the requirements of training equipment, and has better robustness to calculation error and external interference. Next, further research will be conducted on how to ensure the accuracy of the solution in the case that only fewer control points can be obtained under the occlusion and night conditions during the training.

Acknowledgments
This work was financially supported by the Taishan Scholars project of Shandong Province (201511020).

References
[1] Chen L , Armstrong C W , Raftopoulos D D . An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique[J]. Journal of Biomechanics, 1994, 27(4):493-500.

[2] Leung M Y , Lo J , Leung Y Y . Accuracy of Different Modalities to Record Natural Head Position in 3 Dimensions: A Systematic Review[J]. Journal of Oral & Maxillofacial Surgery Official, 2016:2261-2284.
[3] Lepetit V, Moreno-Noguer F, Fua P. EPnP: An Accurate O(n) Solution to the PnP Problem[J]. International Journal of Computer Vision, 2009, 81(2):155-166.

[4] Lucey S, Wang Y, Cox M, et al. Efficient constrained local model fitting for non-rigid face alignment[J]. Image & Vision Computing, 2009, 27(12):1804-1813.

[5] Murphy-Chutorian E, Trivedi M M. Head Pose Estimation in Computer Vision: A Survey[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2009, 31(4):607-626.

[6] Zhou, Erjin, Fan, Haoqiang, Cao, Zhimin. Extensive Facial Landmark Localization with Coarse-to-Fine Convolutional Network Cascade[J]. 2013.

[7] Rocca F, Mancas M, Gosselin B. Head Pose Estimation by Perspective-n-Point Solution Based on 2D Markerless Face Tracking[C] International Conference on Intelligent Technologies for Interactive Entertainment. Springer International Publishing, 2014.

[8] Li S, Xu C, Xie M. A Robust O(n) Solution to the Perspective-n-Point Problem[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012.

[9] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22(11):1330-1334.

[10] Xiong G, Ding T H, Wang P. Perspective-n-point pose measurement with two line array cameras[J]. Optical Engineering, 2013, 52(1):177-182.

[11] Martin Köstinger, et al. Annotated Facial Landmarks in the Wild: A large-scale, real-world database for facial landmark localization[C]// IEEE International Conference on Computer Vision Workshops. IEEE, 2012.