

The Application of Stereo Vision in 3D Modeling of Foot

Ke Zhang^{a*}, Zhao Gao^b

College of Engineering, Zhejiang Normal University, Jinhua, 321004, P.R.China

^ajhkeanu@zjnu.cn, ^bgz@zjnu.cn

Keywords: Shoemaking system, Stereo vision, Epipolar line constraint, Projective rectification, Foot modeling

Abstract. With the development of economy and the improvement of personal consumption, the demands of people for consumer goods are becoming more diversified and individualized. In recent years, the personalized shoemaking system is favored by market, and the traditional shoemaking industry needs new technologies in order to meet people's new demands. The 3D modeling of foot is the foundation to realize individualized shoemaking, meanwhile, the foot parameters corresponding to the 3D foot model is also the important basis of foot re-shaping surgery or correction, therefore, the development and research of new technologies and new equipments of 3D foot modeling have realistic meanings. This paper introduces the application of stereo vision in the foot 3d modeling, firstly the system of stereo vision is set up, and the system is calibrated by using BP Neural Networks. Secondly the foot images are acquired, and the left and right images are corrected in accordance with the principle of epipolar line constraint, the left and right images by projection rectification can finish image matching quickly and accurately, and can construct the 3d foot model. The results of research show that the reconstruction of 3d foot model based on stereo vision is effective.

Introduction

The shape of human foot is complex and diverse, and the surface of foot is typically free-form. Thanks to advances of science and technology, and consumer's increasing demand for personalized products, the more and more free-form surfaces are used in the design of structure and shape of product. The 3D models of human foot have extensive application value and development prospects in the fields of personalized shoemaking and medical plastic or correction. But, being limited to current technological capabilities and cost, the measurement of free-form surface is very difficulty, especially to the soft and sensitive surfaces of human, so it is the urgent demand to research and develop advanced 3D measurement technology of free-form surface for human foot modeling [1].

The 3d measurement of free-form surface can be divided into two broad categories, i.e. contact and non-contact measurement [2]. In the process of contact measurement, the sensing probe must be closely contact with the geometry surface, and then the 3D coordinates of shape surface are obtained after scanned point by point. In the practical application, Coordinates Measurement Machine (CMM) is the typical contact measurement equipment. The non-contact measurement is based on the fundamental principle of optics, acoustics and magnetic, and the 3D coordinates of geometry surface are got from some physical analogs by means of finding effective algorithm, there are no contact between measurement equipments and objects.

The contact and non-contact measurement each has advantages and disadvantages according to different objects. CMM have high accuracy, but its measuring efficiency is low, so it is always used for high precision measurement of conventional parts. However, due to the existing contact pressure on the free-form surface in the process of measurement, there are accordingly a various defeats [3]. On the contrary, the non-contact measurement can overcome shortcomings inherent in the contact measurement, for example, it will not make harm and deformation to some soft and fragile surfaces, especially certain parts of human body, which must be finished comfortably and fast.

Stereo vision is an important branch of computer vision field [4], and it is a typical non-contact measurement, so it is especially suitable for surface measurement of human parts. Stereo vision measurement method has advantages of high efficiency, appropriate precision, simple system structure and low cost, so is very suitable for online and non-contact product testing and quality

control. When the human foot is being measured, the image acquisition can be performed instantaneously, so stereo vision is more advanced measurement method.

In the process of image acquisition, due to existence of various factors of 3D scene that effect image gray scale, such as lighting conditions, the measured object geometry and surface characteristics, noise and the camera features, so it is very difficulty to find the matching relations between corresponding image points, and image matching has been the focus of study in stereo vision field.

To establish the accurate matching between corresponding image points, the important geometrical property of epipolar line constraint of stereo vision system [5] can provide certain facilities for image matching, but, the searching process along such epipolar lines is still very time-consuming and is affected seriously by noise. So, the projective rectification presented by Hartley [6] received enormous attention. After rectification by this method, all epipolar lines, which are on the rectified left and right image planes, are parallel each other and in the horizontal direction, and the conjugated epipolar lines are on the same horizontal line. So that, match searching of corresponding points are reduced from 2D plane to 1D horizontal line, which greatly reduces the difficulty of computing and the possibility of false matches. However, research at home and abroad shows [7] that because all the current algorithms of projective rectification must use complex nonlinear optimization, and the stability and speed of algorithms are still need to be improved and perfected.

In this paper, firstly, the two cameras of stereo vision system are calibrated by using BP neural network. Then the adjustment coefficients of epipolar line relative position are calculated according to the geometric property of epipolar line constraint and the projective rectification formulas is established. The experimental and calculated results indicate that the method of projective rectification presented can avoid complex nonlinear calculation with high stability, and can realize effectively human foot 3D reconstruction.

Fundamental Principle of Stereo Vision System

The fundamental principle of stereo vision is that the same object is observed from two or more viewpoints in order to obtain the instant digital images, then the corresponding images points are matched, and the shape and position of the 3D object surface can be calculated according to the triangulation algorithm [5].

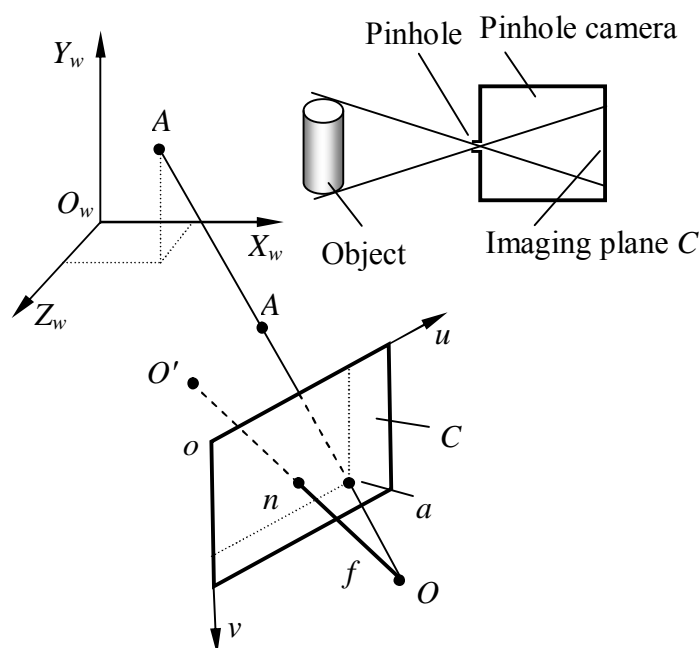


Fig.1 Pinhole imaging model

Pinhole imaging model. The pinhole imaging model is the common ideal linear camera model of stereo vision system, and is the simplification of optical imaging geometry. It can be abstracted as a imaging plane C and a point O outside the plane, i.e. optical center, as shown in Fig. 1.

The 2D coordinate $o-uv$ is set up on the imaging plane C , and the units of u and v is pixel. The perpendicular distance On from optical center O to imaging plane C is called as the focal length of camera f . The connection line AO from the optional point A in the world coordinates to optical center O is called as projection line of point A , and the intersection point a between the projection line and imaging plane C is the image point of A on the imaging plane C , accordingly the optional point A is called as object point. For ease of calculation, the optical center O is placed at back of imaging plane C .

Obviously, the image point in the projection line AO of optional point, such as A' , is also a . Therefore, only depending on the imaging point a on the imaging plane can not determine its corresponding position in the world coordinates. That is to say, the image point a on the imaging plane does not contain depth information, so it is impossible to uniquely determine the object A or A' in the world coordinates by image a .

Stereo vision system. The stereo vision system constructed by two camera of left and right is shown in Fig. 2. For ease of distinction, the corresponding parameters of left and right camera are respectively marked by l and r .

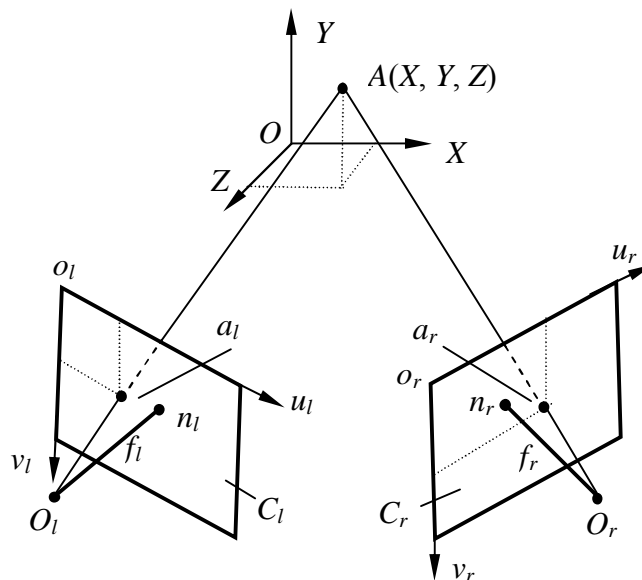


Fig.2 Principle of stereo vision system

The image point of the point $A(X, Y, Z)$ in the world coordinates is respectively a_l and a_r on the imaging plane C_l and C_r . The two image points are the image point of the same object point A , which are called conjugate points. If knowing the two conjugate points in advance, and the projection lines a_lO_l and a_rO_r , which are connection lines from conjugate point to respective optical center O_l and O_r , are constructed, then their intersection point is the object point $A(X, Y, Z)$ in the world coordinates. This is the fundamental principle of stereo vision.

Processing Modules of Stereo Vision System

After construction of stereo vision system composed of two cameras, the subsequent substantive works can be performed one by one, such as image acquisition and processing, image matching, 3D reconstruction and post-processing.

Image acquisition. The acquisition of digital image is the information source of stereo vision, and

the excellent image quality is the guarantee to successfully complete subsequent works. There are several ways to obtain images, which depends on specific occasions and purposes of application. When the stereo images are being acquired, not only the requirements of application must be met, but also the view differences, lighting conditions, camera properties and scene features are all to be considered comprehensively.

System calibration. The camera calibration of stereo vision system is to determine the mapping relation between the coordinates position $a_l(u_l, v_l)$ and $a_r(u_r, v_r)$ of the object point on the left and right imaging planes and the world coordinates $A(X, Y, Z)$, and this is the key step to realize 3D reconstruction of stereo vision. After the calibration, the 3D world coordinates of the object point can be calculated from the 2D image coordinates on the left and right image planes. Among many methods of camera calibration of stereo vision, there is a method that does not need to get every camera physical parameter of stereo vision system, and only some intermediate parameters are need to be calculated, so the process of calculation is simplified, at same time the precision of calibration can be improved to some extent. This method is called implicit camera calibration [8].

Artificial neural network [9] is an implicit camera calibration method, which can perform quickly calibration of stereo vision system. The error back propagation neural network (BP) is currently the most widely used learning algorithm for multilayer forward neural network. When the mapping process of stereo vision system between 2D image coordinates of object points and 3D coordinates of scene is stimulated by the BP neural network, it is not need to build a complex system recovery model, and not need to consider the effects of lens distortion and environment factors, so this can reduce the error as a result of incomplete mathematical mode and improve the reconstruction accuracy of 3D coordinates.

Feature extraction. The image features can be obtained by process of feature extraction, and the properties of image features will directly affect the selection of image matching method. Because of no establishment of generally applicable theory of feature extraction, the matching characteristics is diversified in the research field of stereo vision. Image feature can be pixels or collection of pixels; also can be their abstract expression, such as image structure, image object and relational structure, etc. The common matching characteristics include point feature, line feature and area feature. Generally, image feature of large scale contains more information, and the number of features is small, so the efficiency of matching is high; but the process of feature extraction and feature presentation is difficulty, and the precision of positioning is bad. On contrary, in the case of smaller-scale image feature, the expression and presentation is simply, and the precision of positioning is high; but because of large number of features, and the information is relatively little, so when being matched, more stringent constraints and matching strategy are needed in order to reduce ambiguity and to improve efficiency as possible. Good characteristics of matching should have the ability of separability, invariance, uniqueness and efficiency to address ambiguity.

Image matching. In the stereo vision system as shown in Fig. 2, the image matching is to construct the corresponding relations between the image points $a_l(u_l, v_l)$ and $a_r(u_r, v_r)$ respectively on the imaging planes C_l and C_r of world coordinate object point $A(X, Y, Z)$. The image matching is the most important and difficulty problem of stereo vision, and has been the focus of research of stereo vision. Many factors of the scene, such as lighting conditions, the measured object geometry and surface characteristics, noise and the camera features, are all concentrated in the image grayscale value. Therefore, to accurately match with so many disadvantages, effective constraints are needed to be found.

After obtaining the coefficient correction of projective rectification by using the principle of epipolar line constraint, the original images must be transformed in order to construct the matching relations of image points in the same horizontal line on the transformed left and right imaging planes. In the process of creating new gray images from original gray images, the original image point corresponding to the new image pixel must be found. The process of obtaining corresponding point pixel coordinates (u, v) from new image pixel coordinates (u', v') by means of the reverse relationship was shown as in Fig. 3.

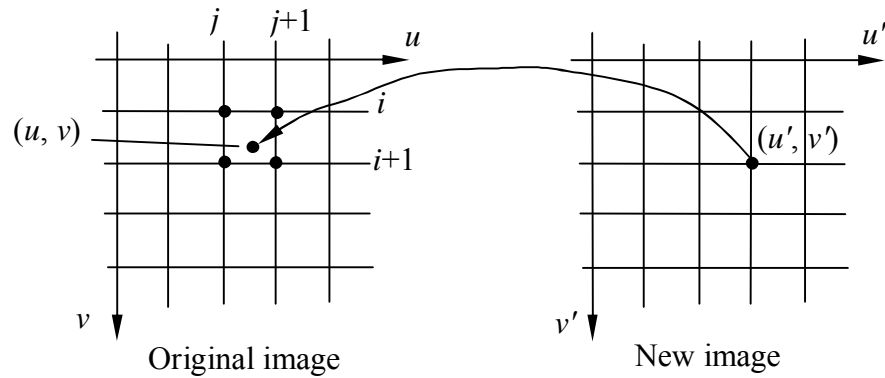


Fig.3 image transformation

3D reconstruction. After finishing the calibration and image matching of stereo vision system, the 3D information reconstruction of surface points of the measured object can be performed by following the trained BP neural network. The factors that affect the precision of 3D measurement include error of camera calibration, digital quantization effects of CCD imaging equipment, feature extraction and precision of match positioning, etc. Therefore, it is needed to reduce or eliminate the adverse effects existing in each phase as much as possible.

Post-processing. The post-processing includes mainly interpolation and texture mapping. Since only the depth information of image points corresponding to image features can be reconstructed when using image characteristics matching algorithm, the interpolation calculation of depth information of surface points is needed in order to perform the 3D reconstruction of the whole visible surface. In the process of interpolation, the most important thing is to protect the discontinuous information of visual surface, so the interpolation action must meet the principle of surface compatibility. The common method of interpolation includes nearest interpolation, bilinear interpolation and spline interpolation.

The texture mapping is to map image texture to the surface of 3D geometry model obtained by reconstruction in order to get true and natural visual effects for geometrical models, and to achieve higher visual requirements.

3D Modeling of Foot

After the match between left and right images, and getting a serious of conjugate feature points, the 3D coordinates of the human foot in the world space can be reconstructed point by point according to the mapping relation between 2D coordinates and 3D world coordinates, so the 3D information reconstruction of serous object point were completed. The realistic 3D models of human foot were obtained from a serious space scattered points through interpolation and texture mapping, as shown in Fig. 4.



Fig.4 3D model of human foot

Conclusion

Considering the human foot characteristics of highly complex and difficult to measure, the non-contact method of stereo vision is adopted. According to the process of stereo vision method, the six phases of work were finished, i.e. image acquisition, system calibration, feature extraction, image match, 3D reconstruction and post-process. The calculation and experimental results show that the stereo vision method can be effective in application of 3D digital modeling of human foot.

References

- [1] Okada S, Imade M and Miyauchi H: Industrial Electronics Society, 1998. IECON '98. Proceedings of the 24th Annual Conference of the IEEE, 1998, pp.1284~1289.
- [2] D.R Shan, Y.L Ke and Y.F Liu: China Mechanical Engineering. Vol. 14, No. 1, (2003), pp.9~12(In Chinese).
- [3] X.Guo, D.M. Peng and J.S. Zu: Aviation Precision Manufacturing Technology. Vol. 40, No. 5, (2004), pp. 10~13(In Chinese).
- [4] J.G. Wang and Y.F. Li: Instrumentation and Measurement Technology Conference, 1999. IMTC/99. Proceedings of the 16th IEEE Volume 2, 24-26 May 1999, 2. pp. 684~689.
- [5] S.D. Ma and Z.Y. Zhang: Computer Vision-Fundamentals of Computational Theory and Algorithms. Science Press. Beijing, China, 1998, pp. 78-80(In Chinese).
- [6] Hartley R. I: International Journal of Computer Vision. Vol. 35, No. 2, (1999), pp. 115~127.
- [7] Loop C. and Z.Y Zhang: Proc. of IEEE conf. on computer vision and pattern recognition. Vol. 1, (1999), pp. 125-131.
- [8] G.Q Wei and S.D Ma: IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(5): 469~480.
- [9] Martin T H, Howard B D and Mark H B: Neural Network Design. Mechanical Industry Press, Beijing, China, 2002, pp. 197-222(In Chinese).
- [10] J.H Zhai, W.X Zhao and X.Z Wang: Journal of Hebei University. Vol. 29, No. 1, (2009), pp. 106-112(In Chinese).