Acta Technica *62* No. 3A/2017, 213–224

Study on binocular stereoscopic vision of Hawk-eye for tennis based on three-dimensional positioning

Cheng Yang[1,2], Li Guo[1]

Abstract. In the game of tennis, there always are some misjudgments or unfair penalties to players. However, "hawk-eye" system, which is brought in, can significantly help referees to make more reasonable penalties, and ensure the matches with justice and equity. A further processing is made to receive the tennis positional data with analyzing the image-forming principle of "hawk-eye" binocular stereoscopic vision, capturing the motion trail of the ball and transmitting to PC end by using wireless sensing. From the perspective of three-dimensional positioning algorithm, non-range-based positioning algorithm is firstly studied to come up with a three-moving beacon positioning algorithm, which is applied to two dimension; directing at three-dimensional space, this algorithm combined with hierarchy coordinated put forward multi-movable beacon hierarchy positioning algorithm: from the perspective of image reconstruction of binocular stereoscopic vision, camera calibration algorithm is studied and a further improvement is put forward. According to a large number of experiments, the proposed positioning algorithm has a higher accuracy. A smaller data size can reduce equipment the consumption. At the same time, the improved method imaging standardization makes a further improvement on acquiring image accuracy.

Key words.
Binocular vision, three-moving beacon positioning, multi-movable beacon hierarchy positioning, camera calibration.

1. Introduction

After a century of development, tennis [1] has become a commercial sport. Tennis players need to gain critical scores as much as possible and make sure they can win victory. However, in some critical balls, the referee may make some controversial penalties. What's more, the spectators would be disappointed by the matches with more unfair penalties. From the perspective of developing tennis and maintaining athletes' rights and interests, it is important to bring "hawk-eye" [2] system into tennis, which can identify the location precisely.

[1]Shijiazhuang Tiedao University, 050043, Hebei, China
[2]Corresponding author; E-mail: chengyang_college@yeah.net

Actually, “hawk-eye” system is an immediate playback system [3]. First, it makes a millimeter partition to the place's three-dimensional space in virtue of PC; then using camera captures the basic data of tennis movement trail from different angles; these data are received by computers to process three-dimensional images; lastly, the images are transmitted to the screen, which can be on the live television. The whole process only takes several seconds.

2. Literature review

As pattern of computer vision, binocular stereoscopic vision [4] can use camera to acquire two located images from different positions. Positioning vanishing point of imaging matching corresponded with computer also can be called anaglyph. From that, the three-dimensional information can be acquired.

Geometric model of camera imaging [5] needs to take intrinsic parameter (corresponding parameter of geometrical characteristic and optical characteristic) and external parameter into consideration. camera standardization [6], and data needs to be collected to acquire the imaging, but the location [7] is the most important step. Data is transmitted to PC end by camera with wireless network. Research on wireless network is always a hot topic, and its related algorithm can be classified into two ones: ranging and non-ranging algorithms. Based on RSS [8] is the most typical algorithm in ranging algorithms. This algorithm can get the node position through the relationship between distance and signal intensity; DV-HOP [9] is one of non-ranging algorithms, which estimates the distance among nodes by using the relationship between cop count and distance.

This paper studies the camera standardization algorithms, and we consider that camera bracket can be adjusted, the semi-auto camera standardization is proposed based on checkerboard plane template. Compared with conventional imaging algorithms, the accuracy of acquiring imaging get further improved. From the viewpoint of 3D, the DV-HOP algorithm is studied and combined with mobile beacon, and a new improved algorithms is proposed.

3. Research methods

3.1. Binocular vision image-forming theory and epipolar geometry

Generally speaking, two cameras are used in binocular vision. As shown in Fig. 1, points $C1, C2$ are respectively the optical centers of the cameras. Point T is the focal point of camera light. Points $(V1, W1)$ and $(V2, W2)$ are the light transmitting nodes $R1$ and $R2$ projected by T on the camera. Points $Q1, Q2$ are the intersections of the vertical direction corresponding to the camera optical center to surface. At the same time, points $(V1, W1)$ and $(V2, W2)$ can also be expressed as the same object T space, namely "conjugate point". Points $C1$ and $C2$ are connected by points $R1$ and $R2$, namely $R1C1$ and $R2C2$ projection lines. They are respectively

extended to point T, which can be represented as coordinates (x, y, Z)

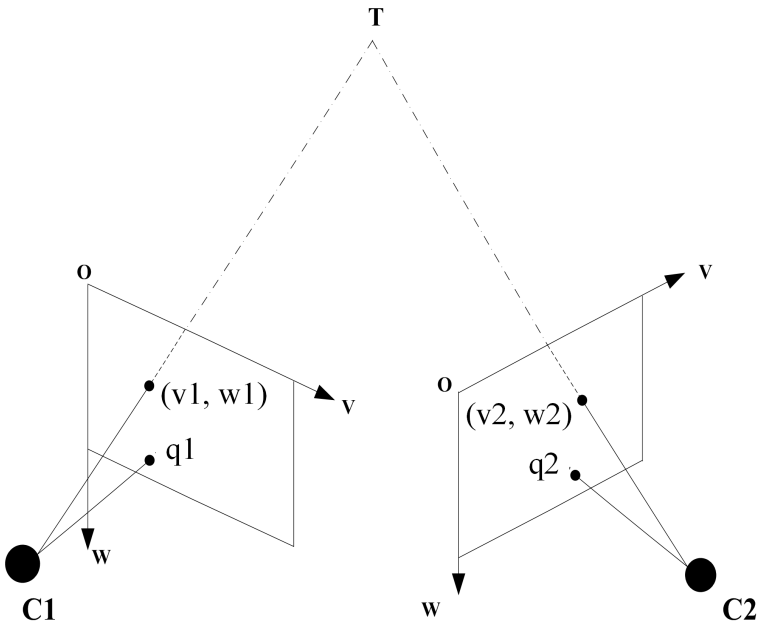

Fig. 1. The Theory of Binocular Stereo Vision

Binocular stereo vision has two structural models: one is that two cameras' optical axis are in a parallel position, which can make the imaging plane formed in the P plane. And if signature is made to camera's photo center, the corresponding baseline is also parallel to the P plane. Optical axis in the plane is perpendicular to the baseline; another is that two cameras are put in a mobile position, but the binocular stereo vision still has some common rules that are unrelated to the cameras' positioning place.

From Fig. 2, it is seen that the left coordinate system of the camera consists of $Q_1A_1B_1C_1$, and right system is composed of $Q_2A_2B_2C_2$. Symbol G in the three-dimensional space corresponds to projection starting point $< g_1, g_2 >$ in the delineation plane of the two cameras. And the two points are extended and intersected at point G in space. A further assumption is put forward. Points Q_1 and Q_2 are very close to their corresponding position. Therefore, Q_1 and Q_2 can be regarded as corresponding photo center nodes. Points C_1 and C_2 have a consistent direction vector with relative space of camera optical axis; G_1 and G_r, respectively, represent the coordinates that intersect with G; a rigid body transformation defined by a translation direction vector Tc and the orthogonal rotation matrix EC represent the geometrical relationship between G_1 and G_r, which is expressed as follows:

$$T_\mathrm{c} = Q_1 - Q_2 \tag{1}$$

$$G_{\mathrm{r}} = EC(G_1 - T_{\mathrm{c}}) \tag{2}$$

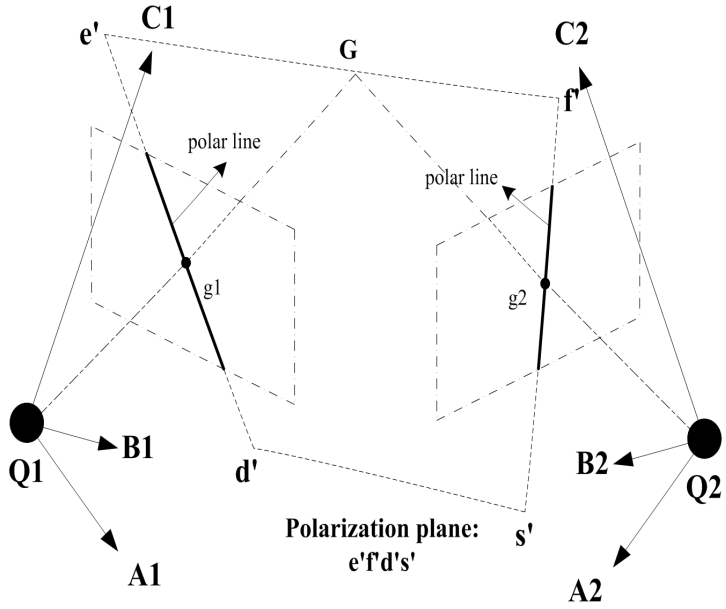

Fig. 2. Epipolar Geometry

Projection point $< g_1, g_2 >$ of two cameras is in $e'f'd's'$ plane with space intersection point, which is also called pole plane. The rest points of the space are in another plane with two projection points as well. At the same time, all the pole planes are collinear, which is called base line. It is the pole ligature. Imaging point of right camera's center Q_2 at the left image is k_l, while center Q_1's image point is k_{r}. So the line is known as $k_l k_{\mathrm{r}}$. When the spatial point G is in an unstable position, the pole plane β rotates with base line, which contributes to a gather of pole plane. In pole plane $e'f'd's'$, camera plane is vertical to optical axis and intersected with two cameras' planes, so w_l and w_{r} in the intersection plane are called polar.

3.2. Improved methods of camera calibration

Camera calibration needs to solve parameters, including internal parameters and external parameters. As shown in Table 1, the internal parameter is the setting parameter of the camera itself, affecting light capture and image focus. It mainly contains central point $< r_x, e_x >$ that is regarded as center for the captured object. However, it is affected by lens distortion and noise. Two focal distances f_s and f_z are involved in the direction of S/Z. The focal length is the distance from the theoretical center to the main focus of the camera lens, which is represented as the ratio of the focal length of the camera to the distance in the horizontal and vertical directions.

At the same time, the non-vertical factor of the camera is set to α. The external parameter is the orientation relation of the camera coordinate system corresponding to the world coordinate system, that is to say, the position relation between the rotation matrix and the translation matrix. In addition, two linear classifications of camera imaging models need to be considered. If it is a nonlinear imaging model, the change parameter of radial lens is $< v1, v2 >$ and the change parameter of the tangential lens is $< t1, t2 >$.

Table 1. Camera parameters

Parameters	Expression	Degrees of freedom
Internal parameter	$V = \begin{bmatrix} f_s & \alpha & r_x \\ 0 & f_z & e_x \\ 0 & 0 & 1 \end{bmatrix}$	5
Distortion (radial direction, tangential direction)	$< v1, v2 >,\ < t1, t2 >$	4
External parameters (rotation matrix, translation matrix)	$R = \begin{bmatrix} rm_1 & . & . & . & rm_m \\ . & . & & & . \\ . & & . & & . \\ . & & & . & . \\ . & . & . & . & rm_n \end{bmatrix},\ T = \begin{bmatrix} l_x \\ l_y \\ l_z \end{bmatrix}$	6

Considering that stand of tennis "hawk-eye" is adjustable, a semi-automatic camera calibration method based on plane template of checkerboard is adopted. A in the template space is assumed to respectively represent the camera coordinate and imaging coordinate of the projection point. In the camera coordinate system, the space coordinate is (X, Y, T, m); in the image coordinate system, the space coordinate of image point a' is (s, z, m). Symbol m is a constant, and the projection relationship between A and a' can be expressed as follows:

$$Wa' = V[RT]A\,. \tag{3}$$

In the above formula, W is a non-zero invariant and external parameter is $[RT]$. Vector $< r_x, e_x >$ serves as the central point, f_s and f_z are the effective focal lengths in the S and Z axes, and α is the non-vertical factor in the S and Z axes, satisfying the condition $\alpha \neq 0$. The intrinsic parameter V has a form of matrix

$$V = \begin{bmatrix} f_s & \alpha & r_x \\ 0 & f_z & e_x \\ 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

It is assumed that the template plane is at world coordinate system $C = 0$; matrix relationship between the template and image is fulfilled. In this formula, mi $(i = 1, 2, 3)$ can express the rotation matrix R in ith row

$$W[s\ z\ m]^{\mathrm{T}} = V[m1m2m3T][XY0m]^{\mathrm{T}} = V[m1m2T][XYm]^{\mathrm{T}} \tag{5}$$

Unit matrix between spatial objects A and a' is expressed as $3{\times}3$ I-matrix, including a nonzero constant factor such that (3) is transmitted into

$$Wa' = IA,\ I = V[m1m2T]\,. \tag{6}$$

The fundamental constraint of camera intrinsic parameter needs to be counted. And I is recorded as $[I1I2I3]$, which is expressed as a product of (6)

$$[I1I2I3] = \theta V[m1m2T]\,. \tag{7}$$

In the expression, θ is a non-zero constant factor. Because the rotation matrix is characterized by orthogonality, $m1$ and $m2$ also meet orthogonality condition. The module value is equal to 1, and then intrinsic parameter of fundamental constraint is

$$\begin{cases} I1^{\mathrm{T}}V^{-\mathrm{T}}V^{-1}I2 = 0, \\ I1^{\mathrm{T}}V^{-\mathrm{T}}V^{-1}I1 = I2^{\mathrm{T}}V^{-\mathrm{T}}V^{-1}I2. \end{cases} \tag{8}$$

Here, $V^{-\mathrm{T}}V^{-1}$ is curve of the second degree, and its corresponding unit matrix has 8 degrees of freedom. In the external parameter, the rotation matrix has three parameters and translation matrix also has three parameters. As a result, in the intrinsic parameter, the matrix has 2 parameters, which means 2 constraint conditions.

3.3. 3.3 Three-mobile beacon positioning algorithm and multi-mobile beacon hierarchy positioning algorithm

Three mobile positioning methods of auxiliary beacon and non-ranging, which is combined mobile beacon with DV-Hop algorithm. Considering characters that beacon produces many virtual beacons in the movement process, one-off positioning is fulfilled. Because it has a smaller consumption and it has a higher positioning precision.

Three-mobile beacon positioning algorithm is a positioning algorithm for topological network and bouncing number accumulation. From Fig. 3, the real distance and route information of every beacon is listed, and $W1$, $W2$ and $W3$ are the beacon nodes. And the three nodes acquire their opposite position information respectively. Bouncing numbers between beacons are gained by unknown node K through collecting virtual data package.

According to distance computation formula, average bouncing distance between $W1$, $W2$ and $W3$ can be worked out. Meanwhile, all bouncing information would

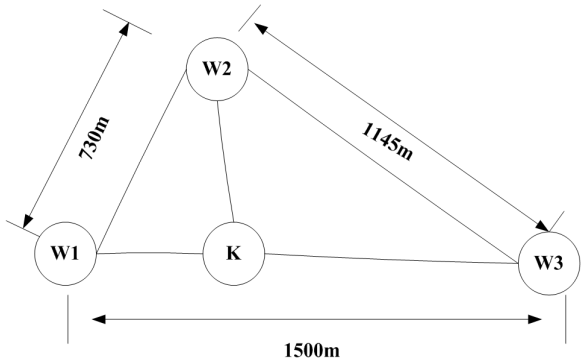

Fig. 3. TMB-DV-Hop Net Work topology

be spread in the whole network.

$$D_i = \frac{\sum \sqrt{(A_i - A_j)^2 + (B_i - B_j)^2}}{\sum f_i},\ i \neq j\,. \tag{9}$$

The distance from the node K to other nodes can be estimated according to the minimum ($Min(\sum f_i)$) and average bouncing distance D_i. This algorithm is an improvement for DV-Hop, and the average distance D_i is worked out unreasonably, especially in the different network environment.

In the two-dimensional environment, with the help of mobile beacon auxiliary, DV-Hop is improved by three-mobile beacon positioning algorithm. However, at most of time, the positioning problem needs to be considered when received data distributed in three-dimensional space. If positioning problem of three-dimensional space is stressed, there will exist a problem—"common plane", which means that four nodes are closely in one plane, and two intersection points resulting from their opposite nodes as radius. Therefore, the unknown node position cannot be estimated. In order to solve this problem, "common plane degree" concept is brought in to expand three-mobile beacon positioning algorithm to three-dimensional space. As a result, multi-mobile beacon hierarchy positioning algorithm is put forward.

Four-side positioning technology is combined with three-dimensional DV-Hop algorithm here. Three-dimensional DV-Hop algorithm also can be divided into three stages:

1) Minimum bouncing number from the beacon nodes is respectively counted by unknown node collection.

2) A minimum bouncing number multiplying average distance is considered as an estimated distance $HopSize_i$ between beacon nodes.

3) When the unknown nodes acquire more than four $HopSize_i$, the space positions of unknown nodes can be worked out.

$$HopSize_i = \frac{\sum \sqrt{(A_i - A_j)^2 + (B_i - B_j)^2 + (C_i - C_j)^2}}{\sum f_i},\ i \neq j\,. \tag{10}$$

In expression (10), $(A_i,\ B_i,\ C_i)$ are the space coordinates of the beacon node i; $(A_j,\ B_j,\ C_j)$ are the space coordinates of beacon node$j$; f_i expresses the bouncing number counted from ith node to other beacon nodes. When distances of four or more unknown nodes are known, a resolution can be carried out by multilateral measurement. If multilateral measurement is changed into four-lateral positioning method, the formulae could be expressed as follows:

$$\begin{cases} (A_1 - A)^2 + (B_1 - B)^2 + (C_1 - C)^2 = L_1^2, \\ (A_2 - A)^2 + (B_2 - B)^2 + (C_2 - C)^2 = L_2^2, \\ (A_3 - A)^2 + (B_3 - B)^2 + (C_3 - C)^2 = L_3^2, \\ (A_4 - A)^2 + (B_4 - B)^2 + (C_4 - C)^2 = L_4^2. \end{cases} \tag{11}$$

In expression (11), L_i is the distance from the unknown node to the beacon node i, which can transform the expression into the form of $XZ = Y$ and we can further get the space vector of (X, Y, Z).

4. Experiment result and analysis

4.1. Standardization simulation and analysis

In the experiment, the adopting chessboard is 10×10 Black and White. There are 81 grids in all except outermost layer. Every grid lens is 2 cm long. The camera is a specialized vidicon for tennis. Out of experimental purpose, narrow baseline configuration is adopted that is 13 cm long. The out image is set up as RGB, its resolution ratio is 640×480 and pixel size is 15(10–3 cm)×15(10–3 cm).

The appointed chessboard is placed in camera shooting scale in front of the camera; the light source is sunlight; a rotational movement is made by control templates. However, the image is not parallel to optical axis. Every time the template is rotated and standardization plank is shot by two cameras, so camera's intrinsic parameter can be counted by the three images. In order to reduce the standardization error, a lot of images should be shot from different angles.

Tool cabinet of MatLab is used to make standardization. In the standardization, the bigger error may occur. At this time, tangential or radial parameters can be imported to reduce the error, but this method has no experience to refer and the judgement is made only by naked eyes. In order not to influence the standardization precision, the tangential or radial parameters are regarded as input value of optimizing standardization algorithm to improve its precision. Table 2 and Table 3 contain the intrinsic parameter between after and before, and the optimized standardization result is more close to actual value.

Adopting default standardization and improved one acquires the corresponding result of Fig. 4 and Fig. 5. The error of extracting image is less than 0.5 pixels; and error range of improved standardization method is more concentrated, so the error is much less.

Table 2. Default internal results of left camera labeling

Parameters	Results
$< fs, fz >$	[485.45754, 483.78545]±[0.94578, 0.93542]
$< rx, ex >$	[375.45754, 342.78545]±[0.97578, 0.88542]
α	0.00000±0.00000
$< v1, v2 >$, $< t1, t2 >$	[–0.02218, 0.02358, 0.00461, –0.00204, 0.00000]
err	[0.13478, 0.0753]

Table 3. Refined internal results of left camera labeling

Parameters	Results
$< fs, fz >$	[484.45754, 482.78545]±[0.94538, 0.93552]
$< rx, ex >$	[375.42554, 340.78545]±[0.95278, 0.89542]
α	0.00000±0.00000
$< v1, v2 >$, $< t1, t2 >$	[–0.02222, 0.02358, 0.00461, –0.00210, 0.00000]
err	[0.12408,0.0621]

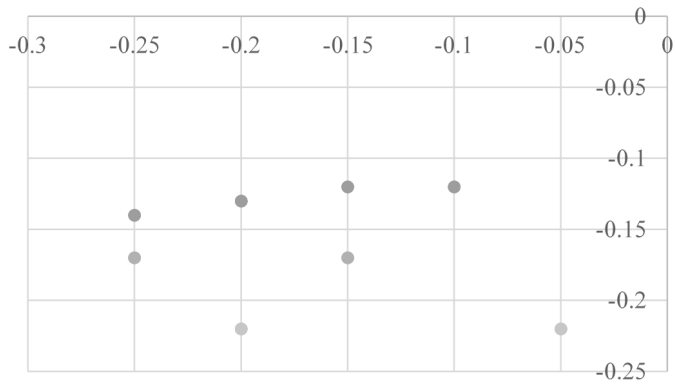

Fig. 4. Projection result 1

4.2. Positioning algorithm simulation and analysis

In every same wireless network (like Fig. 6), the three algorithms have higher positioning precision, which are all less than 0.5. Maximum and minimum average errors of three-mobile beacon positioning algorithms are in lower level, which is better than DV-Hop and MB-DV-Hop. At the same time, DV-Hop precision has

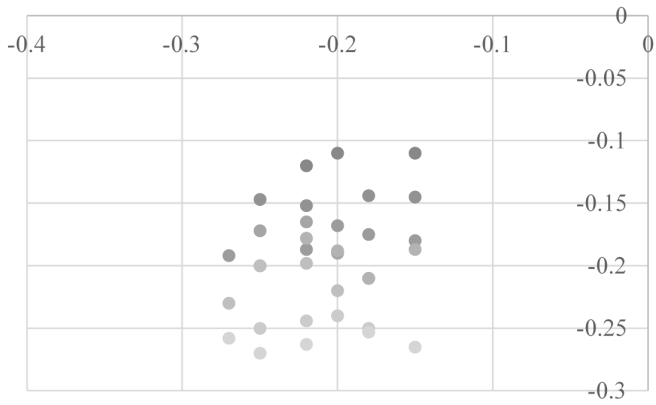

Fig. 5. Projection result 2

a huge fluctuation and its algorithm is not very stable. In the different wireless network (like Fig. 7), TMB-DV-Hop still has a higher precision and is very stable. However, DV-Hop has a huge fluctuation, hardly having any reference.

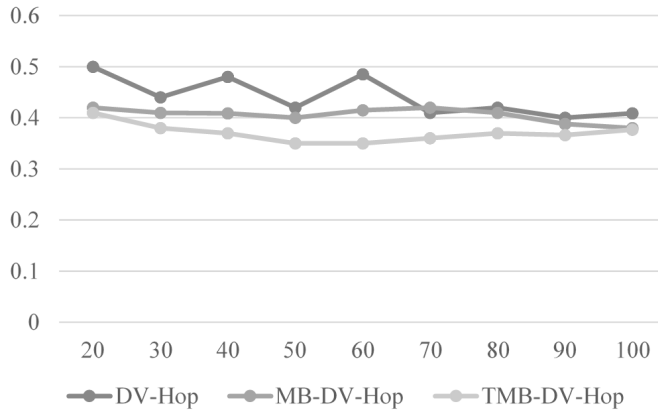

Fig. 6. Precision with XX electricity

DV-Hop in two-dimensional space is extended to three-dimensional space by multi-mobile beacon hierarchy positioning algorithm (like in Fig. 8). Compared with improved algorithm, both improved algorithms and 3D-DV-Hop have lower error. Maximum and minimum errors of multi-mobile beacon hierarchy positioning algorithms are in a lower level. Nodes in 3D-DV-Hop are set up as 10 and 20, and this algorithm precision has a huge fluctuation. As the number of nodes increases, the error is gradually reduced. Generally speaking, however, the average error of multi-mobile beacon hierarchy positioning algorithm is at a lower level and very stable.

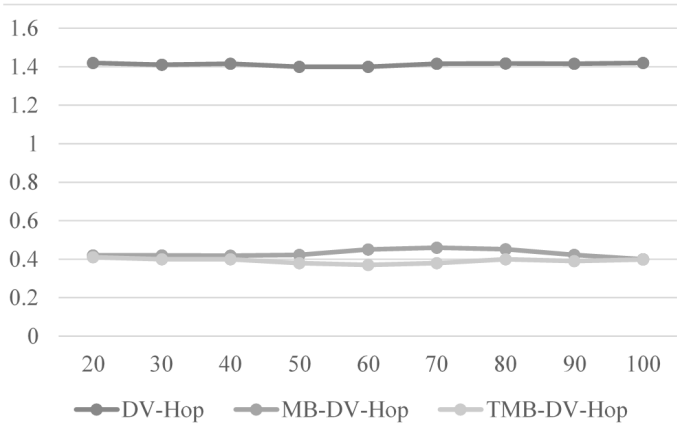

Fig. 7. Precision with XY electricity

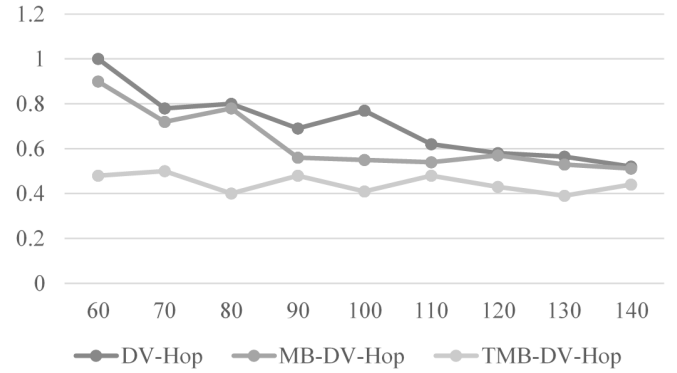

Fig. 8. Precision with 3D space

5. Conclusion

Camera imaging is disposed by binocular stereo vision. Image-forming principle of binocular stereo vision is analyzed, and epipolar geometry is studied from the optic conception. From the hardware, camera's standardization method is improved, which can improve imaging precision further. From two-dimensional plane, DV-Hop is researched and three-mobile beacon positioning algorithm is put forward by virtue of mobile beacon. It is expanded to three-dimensional plane, which results in multi-mobile beacon hierarchy algorithm. A number of experimental results show that when the beacon nodes are constantly changed, the improved algorithm has a higher precision and is in a relatively stable state with smaller data fluctuation.

References

[1] F. Alyas, M. Turner, D. Connell: *MRI findings in the lumbar spines of asymptomatic, adolescent, elite tennis players.* British Journal of Sports Medicine *41* (2007), No. 11, 836–841.

[2] Li Zhao, Ning Xia: *The volleyball movement under the "eagle eye" can see how far–On "eagle eye" effect on the development of volleyball competition system.* Contemporary Sports Technology *4* (2014), No. 20, 141–142.

[3] M. S. Y. Jong, E. T. H. Luk: *Adopting eagle eye in outdoor exploratory learning from the teacher perspective.* IEEE International Conference on Advanced Learning Technologies, 7–10 July 2014, Athens, Greece, IEEE Conference Publications (2014), 617–621.

[4] R. Xiang, H. Jiang, Y. Ying: *Recognition of clustered tomatoes based on binocular stereo vision.* Computers and Electronics in Agriculture *106* (2014), 75–90.

[5] Y. Liu, H. Y. Chen, K. Liang, C. W. Hsu, C. W. Chow, C. H. Yeh: *Visible light communication using receivers of camera image sensor and solar cell.* IEEE Photonics Journal *8* (2016), No. 1, Paper 7800107.

[6] P. Serafinavičius: *Investigation of technical equipment in computer stereo vision: Camera calibration techniques.* Elektronika ir Elektrotechnika *59* (2005), No. 3, 24–27.

[7] H. Zou, X. Lu, H. Jiang, L. Xie: *A fast and precise indoor localization algorithm based on an online sequential extreme learning machine.* Sensors (Basel) *15* (2015), No. 1, 1804–1824.

[8] M. Pajovic, P. Orlik, T. Koike-Akino, K. J. Kim, H. Aikawa, T. Hori: *An unsupervised indoor localization method based on received signal strength (RSS) measurements.* IEEE Global Communications Conference (GLOBECOM), 6–10 Dec. 2015, San Diego, CA, USA, IEEE Conference Publications (2015), 1–6.

[9] S. Kumar, D. K. Lobiyal: *An advanced DV-hop localization algorithm for wireless sensor networks.* Wireless Personal Communications *71* (2013), No. 2, 1365–1385.

Received May 7, 2017