International Journal of Innovative Computing, Information and Control
Volume **10**, Number **2**, April **2014**

ICIC International ©2014 ISSN 1349-4198
pp. 659–669

A NOVEL APPROACH FOR FACE TO CAMERA DISTANCE ESTIMATION BY MONOCULAR VISION

Xiucheng Dong[1], Fan Zhang[1] and Peng Shi[2,3]

[1]Provincial Key Lab on Signal and Information Processing
Xihua University
No. 999, Jinzhou Road, Jinniu District, Chengdu 610039, P. R. China
{ dxc136; ybzf912 }@163.com

[2]College of Automation
Harbin Engineering University
Heilongjiang Province, Harbin 150001, China

[3]College of Engineering and Mathematics
Victoria University
Melbourne, VIC 8001, Australia
peng.shi@vu.edu.au

Received January 2013; revised August 2013

Abstract. *To accomplish the face to camera distance estimation, a method based on monocular vision by using a single camera is proposed. There are three key steps to complete the estimation: extracting and locating the feature region in face, calculating pixel area of characteristic triangle, and constructing the measurement formula. To detect and locate the feature region fast, the traditional AdaBoost algorithm is improved by reducing the features' quantities and expanding samples. The measurement formula which is revealed the relationship between pixel area and distance is derived from the pinhole camera model, camera calibration and area mapping. The measurement formula can be constructed dynamically during the system initialization. The optimum moving range for measurement formula constructing is also defined. After analyzing experimentation, the accuracy of measurement is above 95% and the time each measurement cost is about 230 ms which satisfied the precision and real-time requirement.*
Keywords: Monocular vision measurement, Feature region location, Measurement formula constructing, Image processing

1. **Introduction.** Vision system for distance estimation is extensively concerned as one of the fundamental techniques in the field of machine vision [1]. As shown by increasing interest on the subjects [2,9,20] distance measurement by computer vision is an important component in vehicle guidance, robotic automation and many other computer vision applications. Broadly, the distance estimation algorithm can be divided into 3 categories: stereopsis, motion parallax and monocular cue. By far, stereo vision is the most common study for this problem. Depth-from-motion is based on motion parallax. Both methods need to find correspondence points in multiple images. Assuming that accurate correspondence can be established, both methods can generate very accurate distance estimation. However, the process of searching for image correspondences is computationally expensive and error prone. Many researchers have studied how to use monocular cues for distance estimation. Such studies done both on humans show that cues like texture, texture gradient, linear perspective, defocus and known object size provide information to distance estimation.

In the human and computer interaction system based on computer vision, the distance between user and computer is an essential parameter for HCI algorithm. Besides, it also decides the effect of interaction. In our distance estimation algorithm, we set two constraints: (1) The camera and screen are in the same plane; (2) User's face is straight to the screen (means user's face is parallel to the screen). Under the constraints, it is easy to find that when the user's face is closer to the screen, the image is bigger, the converse is smaller [3]. The distance can be estimated by using the law we mentioned above and the formula we constructed by establishing the mathematic relationship between pixel and distance [4]. The flowchart of distance estimation algorithm is given as Figure 1.

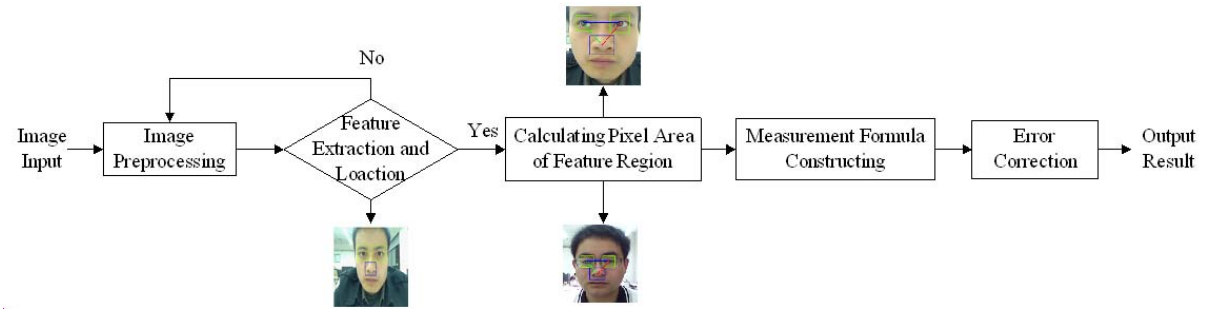

FIGURE 1. Flowchart of distance estimation algorithm

The algorithm is composed of 4 parts:

(1) Image Preprocess: Reducing the resolution of the images and enhancing the image for further processing [5].

(2) Feature Extraction and Location: Extracting the eyes and nose from the face and locating them in the image, constructing the eyes-nose characteristic triangle, calculating the pixel area.

(3) Measurement Formula Constructing: By the use of the optimal movement to construct the mathematic model between pixel and distance, determining the image and distance formula.

(4) Error Correction: Reducing the errors through multiple measurements and constraints.

This paper is organized as follows. Section 2 introduces how to extract and locate the feature region from the face. The principle of distance estimation is derived in Section 3. Experimental results are illustrated in Section 4. Conclusions are drawn in Section 5.

2. Feature Region Extraction and Location in Face.

Extracting and locating feature region in face is the basis of the system. Feature region here means the triangle constructed by eyes and nose. By locating the feature region and constructing the characteristic triangle in image, it is possible to calculate the pixel area of the characteristic triangle. In order to satisfy the real-time requirement, we improved the traditional AdaBoost algorithm, optimized the set of characteristic of eyes and nose and trained a more efficient classifier [6]. The improved AdaBoost algorithm has many benefits such as higher efficiency, stronger robust and higher accuracy. In consideration of different features of eyes and nose, we trained the improved Adaboost algorithm respectively (which is named Eye-AdaBoost and Nose-AdaBoost).

2.1. Feature region extraction.

It is common to use the cascade AdaBoost algorithm to detect the features in face [7]. As for traditional AdaBoost algorithm, the speed of training procedure is slow and the detection relies excessively on the training samples.

In this paper, we fully consider the features of eyes and nose, and improve the speed of training and detection rate by reducing the quantities of features and sample expanding.

In order to improve the detection rate, we use the method of sample expanding to make the features more efficient [8]. The first step is to select a training set of positive and negative samples randomly. After that, training the weak classifier by the set we have selected and combining a collection of weak classification functions to form a stronger classifier. At last, by using the strong classifier to detect the different features, we can expand the characterless samples through detecting the characterless features existing in false detection images. In the same way, we expand the characteristic samples by adding the characteristic features existing in the missing images. Thus, we can have a more efficient strong classifier by re-training the sample. The result of detection is shown as Figure 2.

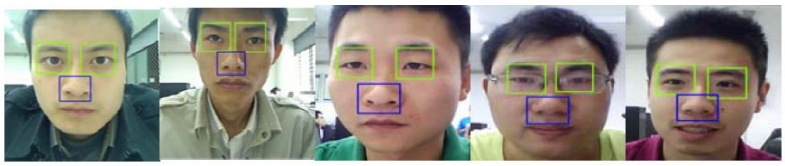

FIGURE 2. Feature region detection

To compare the efficiency between improved and traditional AdaBoost algorithm, we made some experiments. Table 1 shows the results.

TABLE 1. Comparison of feature detection in face

Method	Feature quantities	Detecting time/ms	Accuracy
Traditional AdaBoost (eye detection)	3000	87	97.16%
Eye-AdaBoost	2563	49	98.62%
Traditional AdaBoost (nose detection)	5000	45	98.23%
Nose-AdaBoost	4425	27	98.69%

2.2. **Feature region location.** Locating the pixel coordinate of features in image is the key step before estimation. In order to locate the eyes more accurately, the maximum pixel area for single eye is restricted to 25×25. After the classifier detects the eyes for the first time, the range of searching will be expanded for 5% and will repeat 5 times continually with recording the coordinates. If the eyes can be detected for 5 times, we define that the eyes are detected and the final coordinate is the average value of accumulating 5 searching coordinates. To locate the nose pixel coordinate, we use the Nose-AdaBoost algorithm to detect and record the coordinate directly. The result shows that this method has strong robust and high accuracy [9-12].

3. The Principle of Distance Estimation.

3.1. **Pinhole camera model.** There are four coordinates that have been used: world coordinate, camera coordinate, image coordinate and pixel coordinate [13]. Among the coordinates, the image and pixel coordinates are both in the image plane but the origin and the direction of coordinate axis are different [14]. The image collected by the HD Webcam is projected onto the two-dimension image plane. As for the world coordinate, the pinhole camera model (also known as linear camera model) can satisfy the requirement

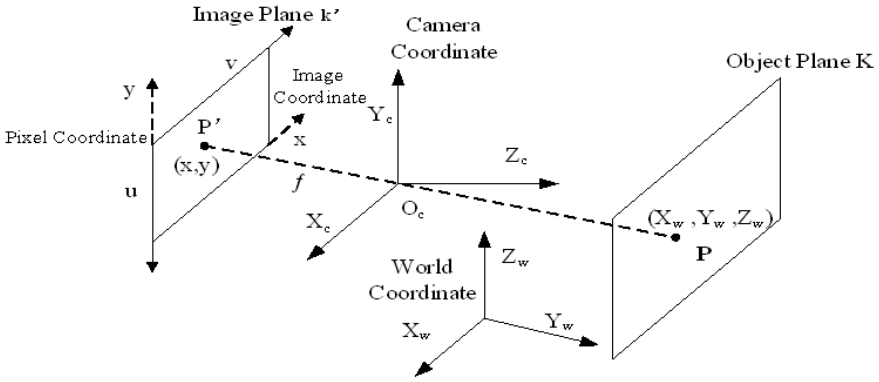

FIGURE 3. Pinhole camera model

[15]. As shown in Figure 3, based on the definition of pinhole camera model, point P on the object plane projects onto the image plane p' though the camera coordinate. According to theory of pinhole camera, we have:

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{1}$$

where, (x, y) is the physical coordinate of p' under the image coordinate, and (X_c, Y_c, Z_c) is the coordinate of P under the camera coordinate. And f is the focus of camera [16].

Images stored in the computer are in the form of digit, we have to create a pixel coordinate that can describe the digital image in the image plane. As shown in Figure 3, (u, v) is the general pixel coordinate we describe above. It is easy to get the following equation by the knowledge of camera calibration:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} d_x & 0 & -u_0 d_x \\ 0 & d_y & -v_0 d_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{2}$$

d_x, d_y describe the pixel counts in unit physical size. To compare the camera coordinate with world coordinate, it is easy to find that:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{3}$$

(X_w, Y_w, Z_w) is the coordinate of point P under world coordinate. R is a orthogonal rotation matrix in the size of 3×3, T is a Translation Vector in the size of 3×1. By substituting Equation (1) and Equation (2) into Equation (3), the relationship between world and image pixel coordinate system can be derived:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_x & 0 & u_0 & 0 \\ 0 & a_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{4}$$

where, $a_x = \frac{f}{d_x}$, $a_y = \frac{f}{d_y}$. Thus, it is available to obtain the object's pixel coordinate from the world coordinate [17-19].

3.2. **Measurement formula constructing.** Measurement Formula Constructing means to establish a formula that can describe the relationship between pixel area and distance that we are willing to estimate. As stated above, the feature region we defined is a characteristic triangle which consists of the iris of the eyes and the nose. Figure 4 shows the characteristic triangle [20,21].

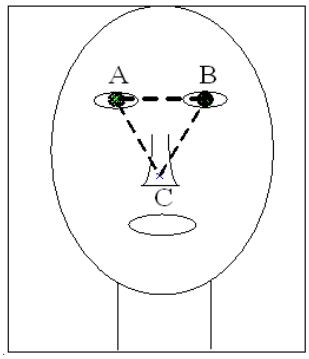

FIGURE 4. Characteristic triangle in face

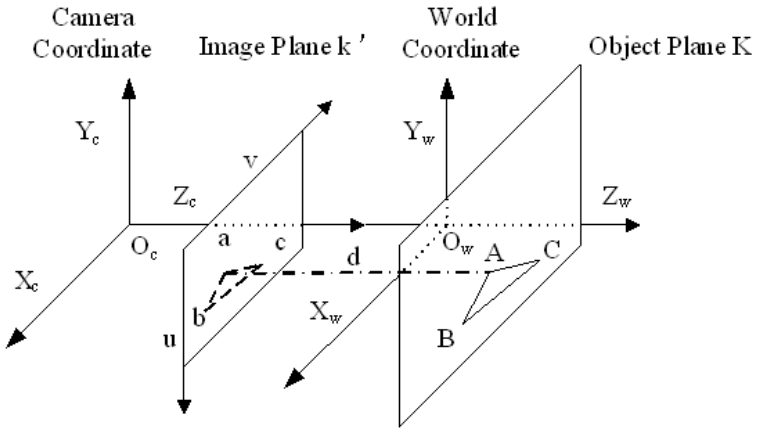

FIGURE 5. Coordinate transformation

To establish the world and camera coordinate as Figure 5 shows, we can obtain: $R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $T = \begin{bmatrix} 0 \\ 0 \\ d \end{bmatrix}$, substituting them into Equation (4), we have:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_x & 0 & u_0 & 0 \\ 0 & a_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} a_x X_w + u_0 Z_w + u_0 d \\ a_y Y_w + u_0 Z_w + v_0 d \\ Z_w + d \end{bmatrix} \quad (5)$$

In this paper, we consider the features in face are in the same plane, so $Z_w = 0$. Further simplifying Equation (5), we can obtain:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{d} \begin{bmatrix} a_x X_w + u_0 d \\ a_y Y_w + v_0 d \\ d \end{bmatrix} \quad (6)$$

$S_{\Delta ABC}$ represents the area of triangle ΔABC in the object plane. Similarly, $S_{\Delta abc}$ represents the pixel area of triangle Δabc in the pixel plane. In the object plane, the

physical coordinate of point A, B, C is (X_A, Y_A), (X_B, Y_B), (X_C, Y_C). In the pixel plane, the pixel coordinate of point a, b, c is (u_a, v_a), (u_b, v_b), (u_c, v_c).

The vector expression for the area of a triangle is $S_{\Delta ABC} = \frac{1}{2}\left|\vec{AB} \times \vec{AC}\right|$ converting it into coordinate expression, we have:

$$\begin{aligned} S_{\Delta ABC} &= \left| \frac{1}{2} \cdot \left| \begin{array}{cc} X_B - X_A & Y_B - Y_A \\ X_C - X_A & Y_C - Y_A \end{array} \right| \right| \\ &= \frac{1}{2} \cdot |(X_B - X_A)(Y_C - Y_A) - (X_C - X_A)(Y_B - Y_A)| \end{aligned} \tag{7}$$

Similarly:

$$S_{\Delta abc} = \frac{1}{2} \cdot |(u_b - u_a)(v_c - v_a) - (u_c - u_a)(v_b - v_a)| \tag{8}$$

Substituting Equation (6) into Equation (8), we can obtain:

$$S_{\Delta abc} = \frac{a_x a_y}{2d^2} \cdot |(X_B - X_A)(Y_C - Y_A) - (X_C - X_A)(Y_B - Y_A)| = \frac{a_x a_y}{2d^2} \cdot S_{\Delta ABC} \tag{9}$$

We can derive the formula between distance and pixel area of characteristic triangle from Equation (9), the formula is given as follows:

$$d = \sqrt{a_x a_y \cdot S_{\Delta ABC}} \cdot S_{\Delta abc}^{-\frac{1}{2}} \tag{10}$$

a_x, a_y are the camera intrinsic parameters that can be calculated by camera calibration. $S_{\Delta ABC}$ is the physical area of characteristic triangle in face. The area of the $S_{\Delta ABC}$ is defined by the volunteer and will not change according to the distance. We define $k = \sqrt{a_x a_y \cdot S_{\Delta ABC}}$ to express the intrinsic coefficient of Equation (10). Thus, we have constructed the measurement formula and derived an ideal formula that can express the relationship between variation of pixel area and distance. We call this formula: Ideal Measurement Formula (IMF).

To establish the measurement formula, we have to calculate the $S_{\Delta ABC}$ first [22,23]. As for every volunteer, measuring the characteristic triangle in face artificially is impossible, that means it have to calculate k automatically in the period of system initialization. In order to solve this problem, we propose a method which utilizes the volunteer's movement vertical to the screen one time in the initialization stage. In this way, it can calculate k automatically and build the measurement formula depending on the specific volunteer. We call this formula Actual Measurement Formula (AMF). The steps to establish the AMF are as follows.

During the system initialization stage, we record the pixel area of characteristic triangle twice. For the first time, it has to make sure the volunteer's face is parallel to the screen (assuming camera and screen are in the same plane), and record the pixel area marked $S_{\Delta abc1}$, so $d_1 = \sqrt{a_x a_y \cdot S_{\Delta ABC}} \cdot S_{\Delta abc1}^{-\frac{1}{2}}$. After that, the volunteer moves toward (or backward) to the screen and record the pixel area $S_{\Delta abc2}$ for the second time, and $d_2 = \sqrt{a_x a_y \cdot S_{\Delta ABC}} \cdot S_{\Delta abc2}^{-\frac{1}{2}}$. The moving distance is marked l, and $l = |d_2 - d_1|$. Substituting d_1 and d_2 into $l = |d_2 - d_1|$, we can calculate $k = \frac{l}{S_{\Delta abc1}^{-0.5} - S_{\Delta abc2}^{-0.5}}$, that means we can build the AMF.

AMF has a great matter with the precision of estimation. In order to match the AMF closer to IMF, we have an experiment to discuss the optimal range of d_1 and l. The intrinsic parameter of the camera that we used in this experiment can be calculated though camera calibration: $a_x = 870.41$, $a_y = 866.26$. The physical area of the characteristic triangle of the volunteer who is selected randomly is $S_{\Delta ABC} = 13.20$ cm^2, substituting it into Equation (10), the IMF is $d = 3154.79 \times S_{\Delta abc}^{-0.5}$. k_l is the intrinsic coefficient of IMF

and k_A is the intrinsic coefficient of AMF. ρ presents the error rate of intrinsic coefficient, $\rho = \left|\frac{k_A - k_l}{k_l}\right| \times 100\%$. Table 2 shows the error rate of intrinsic coefficient in different range of d_1 and l.

TABLE 2. Error rate of intrinsic coefficient for AMF

ρ (d_1 \ l)	5 cm	10 cm	15 cm	20 cm	25 cm	30 cm
20 cm	18.13%	13.79%	9.47%	7.65%	5.51%	3.93%
25 cm	10.10%	0.76%	2.19%	2.46%	1.50%	1.51%
30 cm	8.40%	1.99%	0.52%	0.44%	0.04%	2.01%
35 cm	15.06%	2.08%	1.52%	1.29%	0.72%	2.33%
40 cm	2.29%	2.28%	1.36%	0.01%	2.63%	4.52%
45 cm	3.26%	2.89%	1.42%	0.73%	2.60%	6.57%
50 cm	7.60%	5.22%	6.77%	8.61%	9.17%	13.83%
55 cm	11.19%	10.42%	9.15%	7.73%	9.42%	16.54%

We can learn from the experiment, when $d_1 \in (25, 45)$ and $l \in (10, 25)$, the intrinsic parameter of AMF is closer to the IMF.

4. **Analysis of Distance Estimation.** The experiments are running on a PC (Dominant Frequency: 3.1GHZ, Memory: 2G) with a Logitech HD webcam. The resolution of the input image is 640×480 and the system process 5 frames per second. The interface for measurement is shown in Figure 6.

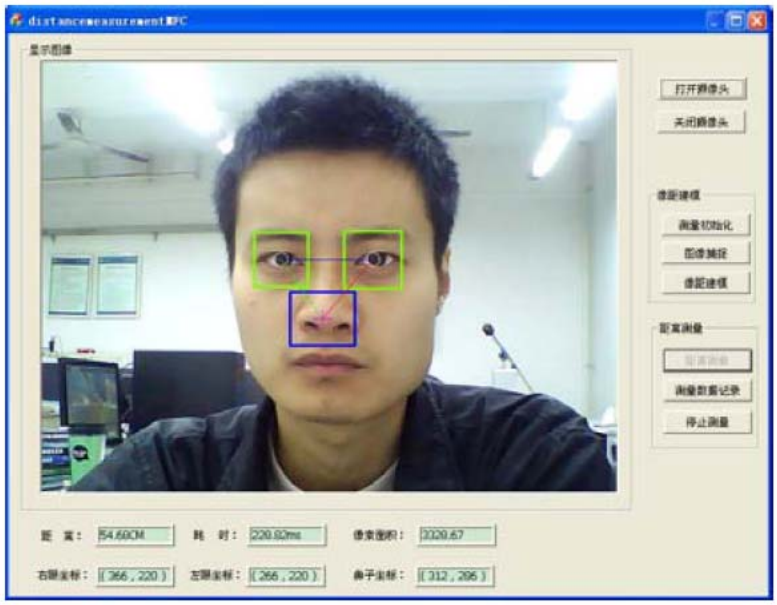

FIGURE 6. Interface for measurement

In order to verify the accuracy, adaptability for the complicated environment and usability of the algorithm, we performed the following experiments:

Test I: In order to verify the usability of measurement formula, we choose a permanent volunteer #1 to measure the distance in the range of 20-80 cm by the step of 5 cm. The IMF of volunteer #1 is $d = 3154.79 \times S_{\Delta abc}^{-0.5}$, the experiment result is given in Table 3.

The verifying experiment indicates that the maximum absolute error of estimation made by IMF is 2.08 cm, and the accuracy is above 97.41%, the feasibility of IMF is verified.

TABLE 3. Distance estimation using IMF

NO.	Actual Distance/CM	Pixel Area	Estimated Distance/CM	Absolute Error/CM	Accuracy	Time Costing/MS
1	20	23729.32	20.48	0.48	97.60%	233.07
2	25	16688.52	24.42	0.58	97.68%	223.45
3	30	11571.31	29.33	0.67	97.76%	229.66
4	35	8422.72	34.38	0.62	98.21%	221.32
5	40	6399.38	39.44	0.56	98.59%	231.95
6	45	5021.33	44.52	0.48	98.93%	236.62
7	50	4062.81	49.49	0.51	98.99%	229.32
8	55	3330.17	54.67	0.33	99.40%	239.46
9	60	2825.13	59.35	0.65	98.92%	234.96
10	65	2452.33	63.71	1.29	98.01%	242.82
11	70	2119.83	68.52	1.48	97.89%	233.56
12	75	1694.51	76.64	1.64	97.82%	231.32
13	80	1639.07	77.92	2.08	97.41%	242.17

TABLE 4. Distance estimation using AMF

NO.	Actual Distance/CM	Pixel Area	Estimated Distance/CM	Absolute Error/CM	Error %	Time Costing/MS
1	20	23618.33	20.83	0.83	4.15%	221.36
2	25	15363.17	25.78	0.78	3.12%	231.32
3	30	10980.5	30.68	0.68	2.27%	224.62
4	35	8120.17	35.50	0.50	1.43%	221.79
5	40	6448.27	39.52	0.48	1.20%	231.33
6	45	4980.17	45.37	0.37	0.82%	234.98
7	50	3970.33	50.81	0.81	1.62%	229.33
8	55	3294.67	55.73	0.73	1.33%	239.64
9	60	2773.33	60.83	0.83	1.38%	220.63
10	65	2378.67	65.97	0.97	1.49%	245.19
11	70	2183.17	68.64	1.36	1.94%	238.61
12	75	1831.83	73.94	1.06	1.41%	240.31
13	80	1660.67	78.29	2.14	2.68%	226.65

In addition, in order to verify the reliability of AMF which is established dynamically during the system initialization, we have the same experiment. The AMF of volunteer #1 is $d = 3239.29 \times S_{\Delta abc}^{-0.5}$, the error rate of intrinsic coefficient is $\rho = 1.52\%$. The result is shown in Table 4.

By analyzing the data, we can easily find that the absolute error is less that 2.14 cm and the maximum error rate is 4.15%. Considering the absolute error and error rate are acceptable, we can conclude that the estimation made by AMF is feasible.

Figure 7 shows that when the error rate of intrinsic coefficient ρ is less than 3%, the accuracy of measurement is higher that 95%. And the higher the ρ is, the lower the accuracy will be.

Test II: In order to test the system's adaptability in different illuminations, we also choose volunteer #1 to make the experiment under 4 different illuminations. ρ represents the accuracy of measurement.

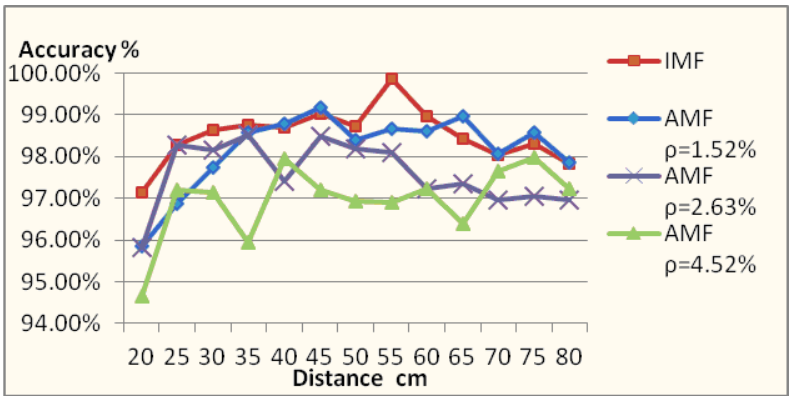

FIGURE 7. Accuracy comparison in different ρ

TABLE 5. Result comparison under 4 different illuminations

Condition	Accuracy in specified distance ρ (%)			Effective Range /cm ($\rho \geq 95\%$)	Time Costing
	30 cm	55 cm	70 cm		
Condition (1)	97.54%	99.31%	97.32%	18-86	233 ms
Condition (2)	97.11%	98.92%	97.04%	21-81	241 ms
Condition (3)	95.12%	96.45%	93.23%	27-66	272 ms
Condition (4)	96.29%	97.94%	95.21%	24-73	253 ms

Condition 1: Experiment is run in daytime. There are 2 incandescent lamps above the volunteer without intense light source disturbing. The illumination is adequate during the whole experiment.

Condition 2: Experiment is run in daytime. There is no incandescent lamp above the volunteer without intense light source disturbing. The illumination is medium during the whole experiment.

Condition 3: Experiment is run in daytime. There are 2 incandescent lamps above the volunteer with intense light source disturbing. The illumination is strong during the whole experiment.

Condition 4: Experiment is run in night. There are 2 incandescent lamps above the volunteer without intense light source disturbing. The illumination is weak during the whole experiment.

The measurement result is given in Table 5.

Learning from Test II we can conclude that the system performs stable under different illuminations. The final result is reliable, but the accuracy of measurement is changed great according to the illuminations. The system performs best in the condition of adequate illumination. When the system is disturbed by the outside light source, the accuracy and effective range is decreased and real-time performs worse simultaneously.

Test III: In order to verify the system's adaptability for different users, we choose 6 volunteers marked #1-#6 to test the measurement accuracy in range of 20-80 cm.

To make the test realistic, we constrain all the tests are running in Condition (1) which is indicated in Test II. The former 3 volunteers do not wear glasses while the latter 3 do. The result is shown in Figure 8.

The system can detect features and accomplish the distance estimation for all the volunteers. The usability of the system is good. After analyzing the result, we can find

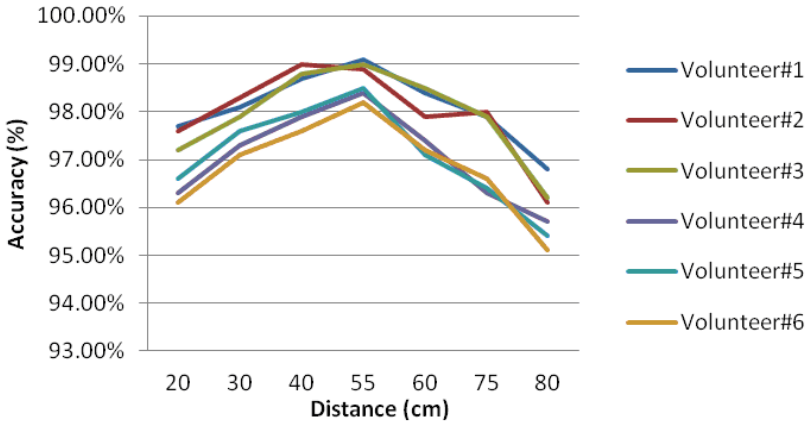

FIGURE 8. Accuracy between different volunteers

that the accuracy which estimated by the volunteers who do not wear glass is higher than those who wear. However, for all the volunteers, the accuracies are above 95%.

5. **Conclusions.** In this paper, we propose a novel approach for face to camera distance estimation by monocular vision. The estimation system has a simple structure and can get a good accuracy. With the Eye-AdaBoost and Nose-AdaBoost algorithm put forth in this paper, the system can quickly extract the characteristic area of face and locate the coordinate of eyes and nose and construct the characteristic triangle in pixel coordinate. Based on the pinhole camera model and coordinate transformation, we can deduce the Ideal Measurement Formula. To establish the AMF, we design a new method which utilizes the volunteer's movement vertical to the screen one time during the system initialization stage. The specific distance and system adaptive experiments verify the feasibility of proposed method and the general applicability of the system. The experiment results show that system's performance is robust under complicated environment, the accuracy is high in the effective range and the system operating frequency is about 5 frames per second.

Acknowledgment. This work was partially supported by the Ministry of Education of People's Republic of China, ChunHui Project (Z2012028), the National Key Basic Research Program (973), China (2011CB710706, 2012CB215202), and the 111 Project (B12018).

REFERENCES

[1] G. Huang, G. Li and B. Wang, Evolution for monocular vision measurement, *Acta Metrologica Sinica*, vol.25, no.4, pp.314-317, 2004.
[2] C. Hsu, M. Lu and W. Wang, Distance measurement based on pixel variation of CCD images, *ISA Transactions*, vol.48, pp.389-395, 2009.
[3] C. Li, X. Jia, H. Li, L. Deng and X. Shi, Digital image processing technology applied in level measurement and control system, *Procedia Engineering*, vol.24, pp.226-231, 2011.
[4] Z. Jing, F. Xue and Z. Li, Space target location measurement based on monocular vision, *Transducer and Microsystem Technologies*, vol.30, no.3, pp.125-127, 2011.
[5] W. Chen, B. Liao, X. Xu et al., Piecewise histogram equalization based image enhancement, *Journal of Communication*, vol.32, pp.153-158, 2011.
[6] P. Viola and M. Jones, Robust real-time face detection, *International Journal of Computer Vision*, vol.57, no.2, pp.137-154, 2004.
[7] W. Wang, F. Huang and J. Li, Face detection and recognition by LBP pyramid, *Journal of Computer-Aided Design & Computer Graphic*, vol.21, pp.94-100, 2010.

[8] Y. Zhang and P. He, A revised AdaBoost algorithm – M-Asy AdaBoost, *Transaction of Beijing Institute of Technology*, vol.31, no.1, pp.64-68, 2011.

[9] L. Gan, J. Zhu and D. Miao, Application of expansion Haar feature in eye detection, *Journal of University of Electronic Science and Technology of China*, vol.39, no.2, pp.247-250, 2010.

[10] Z. Du and Y. Wang, Eye locating algorithm for frontal view face image, *Journal of Computer-Aided Design & Computer Graphic*, vol.21, no.6, pp.763-769, 2009.

[11] J. Ai, D. Yao and Y. Guo, Eye location based on blocks, *Journal of Image and Graphic*, vol.12, no.10, pp.1841-1844, 2007.

[12] M. Hao and X. Dong, Precise human eye pupil localization algorithm, *Computer Engineering*, vol.38, no.8, pp.141-143, 2012,

[13] B. Li, D. Xu, S. Feng and A. Wu, Visual perception theory guided image perceptual depth estimation, *International Journal of Innovative Computing, Information and Control*, vol.3, no.6(B), pp.1625-1634, 2007.

[14] M. Lu, W. Wang and C. Chu, Image-based distance and area measuring systems, *IEEE Sensors Journal*, vol.6, no.2, pp.495-503, 2006.

[15] Y. Han, Z. Zhang and M. Dai, Monocular vision system for distance measurement based on feature points, *Optics and Precision Engineering*, vol.19, no.5, pp.1111-1117, 2011.

[16] Z. Zhang, A flexible new technique for camera calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.22, no.11, pp.1330-1334, 2000.

[17] M. C. Lu, W. Y. Wang and C. Y. Chu, Image-based distance and area measuring systems, *IEEE Sensors Journal*, vol.6, pp.495-503, 2006.

[18] A. Carullo, F. Ferraris, S. Graziani, U. Grimaldi and M. Parvis, Ultrasonic distance sensor improvement using a two-level neural network, *IEEE Transactions on Instrumentation and Measurement*, vol.45, no.2, pp.677-682, 1996.

[19] M. De Santo, C. Liguori, A. Paolillo and A. Pietrosanto, Standard uncertainty evaluation in image-based measurements, *Measurement*, vol.36, no.3-4, pp.347-358, 2004.

[20] J. Phelawan, P. Kittisut and N. Pornsuwancharoen, A new technique for distance measurement of between vehicles to vehicles by plate car using image processing, *Procedia Engineering*, vol.32, pp.348-353, 2012.

[21] T.-H. Wang, M.-C. Lu, C.-C. Hsu, C.-C. Chen and J.-D. Tan, Liquid-level measurement using a single digital camera, *Measurement*, vol.42, no.4, pp.604-610, 2009.

[22] Y. Wei, Z. Dong and C. Wu, Depth measurement using single camera with fixed camera parameters, *IET Computer Vision*, vol.6, no.1, pp.29-39, 2012.

[23] X. Baro, S. Escalera, J. Vitria, O. Pujol and P. Radeva, Traffic sign recognition using evolutionary Adaboost detection and forest-ECOC classification, *IEEE Transactions on Intelligent Transportation Systems*, vol.10, no.1, pp.113-126, 2009.