

Article

Instant Preview and Analysis System Implementation of Intelligent Stereo Cameras

Hsing-Cheng Yu ^{1,*}, Yuan-Zhi Tsai ¹ and Hsieh-Hung Tsai ²

¹ Department of Systems Engineering and Naval Architecture, National Taiwan Ocean University, 2 Pei-Ning Road, Keelung 20224, Taiwan; hcyu@ntou.edu.tw (H.Y.); a0970660117@gmail.com (Y.T.)

² Electronics and Optoelectronics Research Laboratories, Industrial Technology Research Institute, 195, Sec. 4, Chung Hsing Rd., Chutung, Hsinchu 31040, Taiwan; jack_tsai@itri.org.tw

* Correspondence: hcyu@ntou.edu.tw; Tel.: +886-2-24622192 ext. 6059

Abstract: This study presents an instant preview and analysis system implementation of intelligent stereo cameras (ISCs). A parameter optimization prototype adopted the instant preview and analysis system of the ISCs has been achieved the automatic alignment function, and obtained optimal stereo films with the automatic alignment function by adjusting gap and angle between dual cameras. The instant preview and analysis system of the ISCs with parameter optimization can enhance the quality of stereo films effectively and reduce filmed errors and save retouched cost and time in harsh filmed environment.

Keywords: auto-alignment; intelligent stereo camera; stereo film; three-dimensional

1. Introduction

One of the main challenges in three-dimensional (3D) displays is how to supply high quality of stereo video, and it may be an obstacle for developing in three-dimensional. Nowadays, it depends on other countries for advancing equipment that including software and hardware for making stereo movies. The equipment is not only expensive but also need plenty of training time for train new employees. To solve those problems, this study develops an instant preview and analysis system implementation of intelligent stereo cameras (ISCs). It can support the filmmakers efficiently not only to reduce making film time and cost efficiently, but also can obtain stereo films of higher quality.

Several design concepts getting better stereo films are presented in literatures. First, optimal parameters of stereo cameras are presented. In 1986, parameters of stereo cameras could be obtained by solved the object of feature position and relationship in space [1]. In 1999, in order to solve correcting object in wide range of shooting space, it became to finish the process of correct for viewing many images as feature points in any positions in the same time [2]. In 2004, the way of correct stereo cameras was used to three feature points in a line for correct object, it can simply the process of stereo cameras as well [3]. In 2010, the close-form solutions of parameters could be solved for correcting stereo cameras, and it could finish correcting stereo cameras just only need to one image [4].

Second, graphical user interfaces of stereo cameras are proposed in many studies. In 2003, the parallax problems could be solved by tracking feature points interior [5]. In 2004, according to feature spots in stereo images, distortion of lens and luminance could be improved [6]. In 2005, in terms of the position of 3D and motion of object, the problems of stereo camera positions and reflections could be adjusted by drawing software [7].

Third, automatic controllers and driving motor equipment in stereo cameras are offered, and stereo films can be corrected automatically after setting stereo cameras on robots. The robots could

get useful signals and information for correct and position [8]. In 2005, it was used to neural network for correct stereo cameras and aligned the plane position of dual visual system [9].

In summary, this study proposes an instant preview and analysis system implementation of the ISCs, and an automatic alignment function can be achieved by doubling two images of cameras and the ISCs can assist filmmakers to avoid filmed mistake probability before making stereo films. During making stereo films, it can capture continually stereo video streams from controlled depth of field in real time, can draw gray scale film in depth immediately, and can detect video information whether mistakes or not.

2. Instant Preview and Analysis System of Intelligent Stereo Cameras

The instant preview and analysis system of the ISCs is utilized in this study to improve the process and efficiency on the stereo cameras. The system of the ISCs includes equipment hardware and imagination software as shown in Figure 1, and it can be divided into four architectures, i.e., stereo vision camera and equipment, instant preview and replay function, stereo imagination analysis system, and programmable camera stand, respectively. The ISCs can finish stereo film works effectively.



Figure 1. Instant preview and analysis system of the ISCs.

3. Stereo Theory

3.1. Stereo Vision Camera and Equipment

The 3D visual images are generated by the parallax effect which is called binocular parallax. Binocular parallax depends on the different perspectives of the eyes, and dual cameras work as eyes for 3D visual phenomenon. The cameras can get higher 3D visual phenomenon when the dual cameras get closer. The dual cameras have to be corrected the position before the cameras make stereo films. According to the calibration method for positioning the stereo cameras, a schematic diagram of the dual cameras in the instant preview and analysis system of the ISCs can be illustrated in Figure 2.

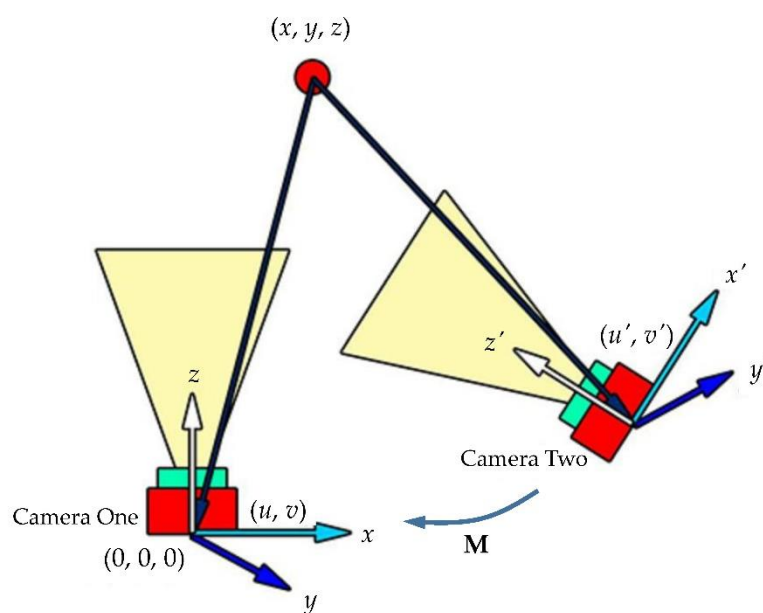


Figure 2. A schematic diagram of the dual cameras in the instant preview and analysis system of the ISCs.

The six degree of freedoms of the dual cameras, e.g. linear motion and rotation in x - y - z axes, always need to be adjusted. The relationship between the dual cameras can be expressed in

$$\mathbf{M}(x \ y \ z \ 1)\mathbf{T} = (x' \ y' \ z' \ 1)\mathbf{T} \quad (1)$$

where \mathbf{M} is the matrix relationship between the dual cameras; x, y, z are the position of the camera one, respectively; x', y', z' are the position of the camera two, respectively. The \mathbf{M} is in form of

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

where m_{11} , m_{12} , m_{13} , m_{21} , m_{22} , m_{23} , m_{31} , m_{32} , m_{33} are the rotation variables, respectively; and m_{14} , m_{24} , m_{34} are the linear motion variables respectively.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \mathbf{M} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} \cdot x + m_{12} \cdot y + m_{13} \cdot z + m_{14} \\ m_{21} \cdot x + m_{22} \cdot y + m_{23} \cdot z + m_{24} \\ m_{31} \cdot x + m_{32} \cdot y + m_{33} \cdot z + m_{34} \\ 1 \end{bmatrix} \quad (3)$$

To simplify the formula, the image projection on another camera, i.e., u' and v' , can be expressed by

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} \frac{x'}{z'} \\ \frac{y'}{z'} \end{bmatrix} = \begin{bmatrix} \frac{x + m_{14}}{z} \\ \frac{x + m_{24}}{z} \end{bmatrix} \quad (4)$$

The dimensions should be adjusted to overlap two images of the dual cameras. The dual cameras also need to synchronize the time code of the cameras because the time code is utilized to match voice, images, and special effects for making films. There are four data in the time code, which includes synchronize, parameters of variables, time, and user information.

3.2. Instant Preview and Replay Function

The users can set mode for stereo video depend on different situation in a suitable video situation. The software can count the parallax of the dual cameras automatically. After the software count, the software would show the date of parameter optimization immediately. The cameras have three operational modes which are landscape mode, dynamic portrait mode, and multi-dynamic portrait mode.

The analysis system of stereo imagination depends on the technique of feedback streaming data in stereo images. The ISCs uses a personal computer (PC) to check stereo signals by two signal imaginations and makes the dual cameras synchronize. The analysis system of stereo imaginations can be classified into two classifications which are analysis of stereo streams and analysis of human factors.

3.2.1. Analysis of Stereo Streams

Analysis of stereo streams can detect uncomfortable feeling from stereo images, and the stereo streams can analyze the light, depth of field, reflective, flash, and pollution. After analyzing of stereo streams, it would show the messages about errors, as shown in [Figure 3](#).

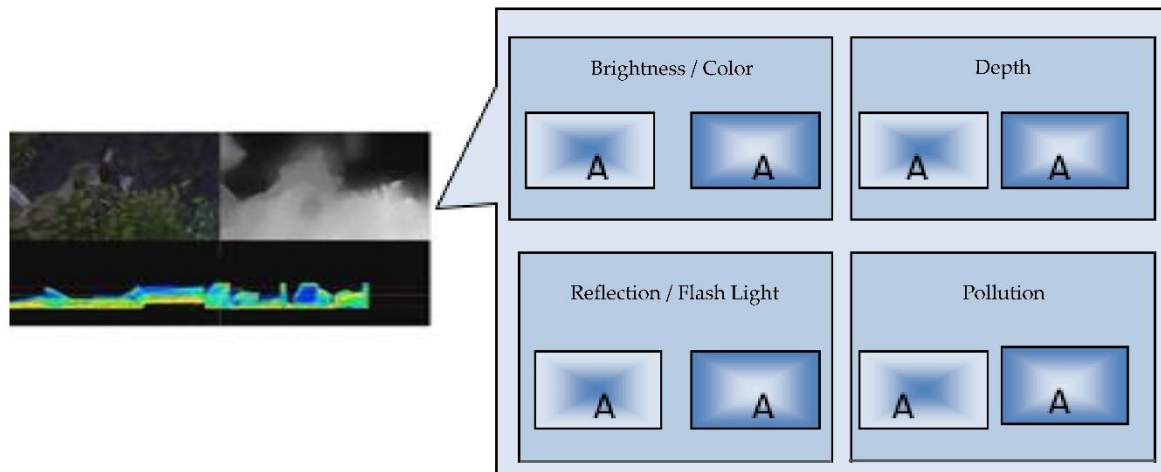


Figure 3. Analysis of stereo streams.

3.2.2. Analysis of Human Factors

In order to analyze human factors, the analysis of human factors depends on equipment to analyze the signals of human eyes and varieties of eyes. After analyzing the human factors, it can get the data about human factors of stereo films. There are three important parameters, i.e. parallax, angles, and speed of movement, to capture users feeling for watching stereo films.

3.3. Programmable Camera Stand

In order to control the dual cameras, the software needs to adjust the angles of the dual cameras, i.e. height, tilt, yaw, and roll, as shown in Figure 4.

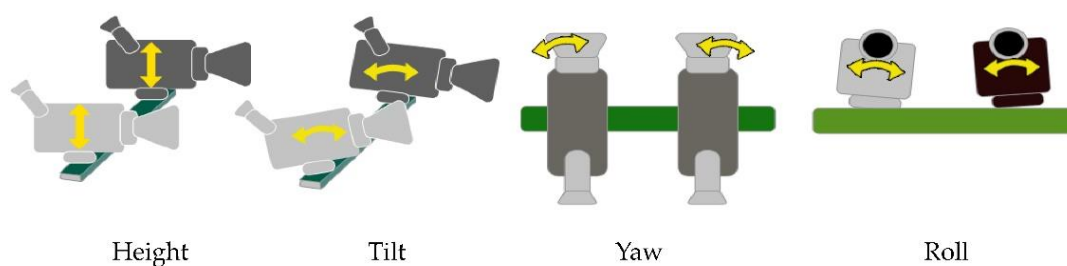


Figure 4. Programmable parameters of the stereo cameras.

To simplify the control parameters, the stereo cameras need to get the information through the stereo cameras. Next, the PC outputs the control of interface, lens, and correction to adjust the camera mechanism by controlling motors. Otherwise, after the mechanism was controlled by the motors, the motors need to be recorded and debugged by feedback signals in the PC, as shown in Figure 5.

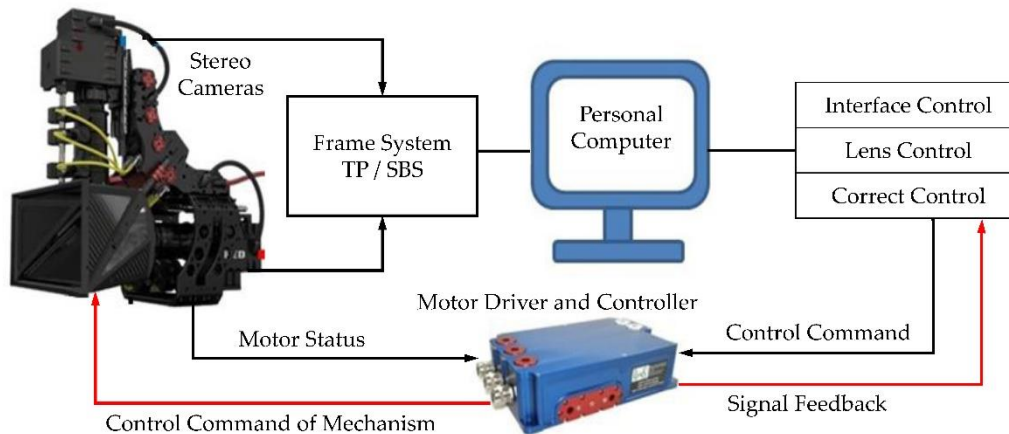


Figure 5. Programmable controller system.

4. Experimental Results

4.1. Optical Axis Correction Architecture of Stereo Cameras

To get a higher quality for correcting the optical axis of the stereo cameras, the ISCs improve the correction method in cradle hand by hand. The ISCs make cradle head holding still and adjusting the feature points to correspond central of images. It can correct the optical axis efficiently by semi-automatic software. The correction method puts a big screen as a big monitor and connects to the PC. After the PC obtains the images, the big monitor would connect to PC monitor and output a correcting image to regulate the optical axis, as shown in [Figure 6](#).

After the software detects the correcting image, the image would show on the monitor and make the image on the central of the monitor. The software would divide into two color which are yellow and blue regions. The two regions can get a cross line and delete the interference on background. After the software gets a cross line, the ISCs can get central position of images by using Hough transform method.

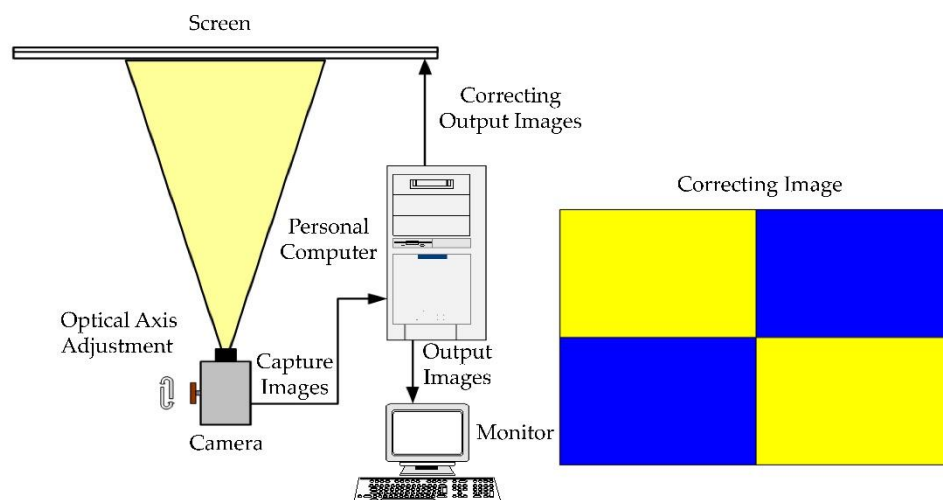


Figure 6. Optical axis correction architecture of the stereo cameras.

4.2. Analysis of 3D Effect on Stereo Images

(1.) Detect Color Contrast

The software can detect two pictures and show analysis results for users. The users can set high contrast parameters to avoid moving blur.

(2.) Auxiliary Correct Two Images

The software can show the results of analysis and green marks. The software also shows the information for users to set parameter efficiency.

(3.) Information of Double Two Images

The function is utilized to conform the parallax of eyes, and it can set the mode of single images, stereo images, and edges.

(4.) Ghosting Analysis

The function of ghosting analysis is divided the images, and compared different colors in the same region. If there are too many different colors in the same region, it would show a warning signal, as shown in [Figure 7](#).

(5.) Stereo Window Violation Analysis

In order to detect the analysis of the stereo window violation, it only needs to judge whether parallax or not near the stereo window violation. After the software detected the analysis of the stereo window violation, it needs to divide the stereo window violation into many varied regions in statistics.

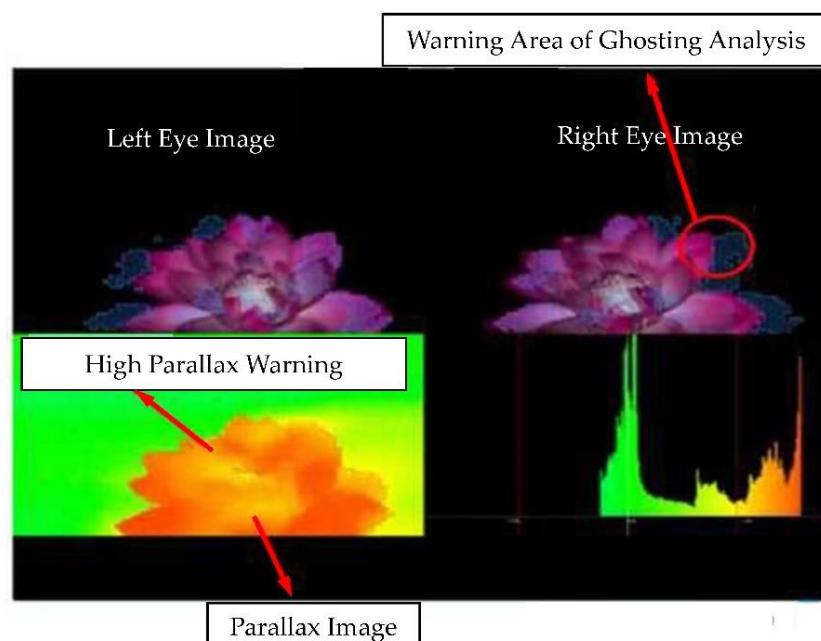


Figure 7. Warning signal of the ghosting analysis.

4.3. Analysis of Brightness Loss

The brightness loss has higher impact on the brightness of screen images between 2D and 3D modes. Otherwise, the brightness loss on the screen images has a significant decrement depended on the change from 2D to 3D mode. The analysis of brightness loss is explained and analyzed respectively below.

(1.) Scene Recording Brightness

Two lighting groups with 3200 and 7500 lumens (lx) are used as the scene recoding brightness. Moreover, the scene recoding brightness shows the two variables in reasonable consideration in statistical results.

(2.) Interaction Values of the Variables

It makes greater or lesser degree with each variables interacts with the others variables, but a given variable has lesser impact on the screen brightness than a major independent variable.

(3.) Screen Size

The screen size is the most important environmental variable when measuring brightness in a darkroom. The model of linear regression model assumes three screen sizes as separate dummy variables. The statistical results show an important impact on the brightness, and larger screens have a positive effect on the value of brightness, falling within the range of reasonable consideration. Therefore, the interactions and screen sizes between other factors importantly affect the brightness.

(4.) Exposure Value of Photographic Equipment

In the experiment with different shutter and aperture conditions, the setup has a direct effect upon the image brightness. A larger exposure value (EV) means less exposure, and the statistical results are consistent with the finding as well.

(5.) 2D or 3D modes

The variable is the major consideration in the experiment, and its important is apparent in the statistical analysis results. Hence, a statistical coefficient impacts on the screen image brightness essentially. The linear regression model and the experimental measurement prove that the 2D image brightness is higher than that of the 3D images. To estimate the value of screen image brightness, the linear regression sample takes the 2D and 3D samples as dummy variables. The expected screen display mode is set in the linear regression model, so that the estimated image brightness value can be obtained.

4.4. An Image Brightness Regression Model

The model is used to analyze image brightness of the screen relationship with the following variables:

- Independent variables (X): (1.) Screen size; (2.) Screen recording brightness; (3.) 2D or 3D mode; (4.) EV of photographic equipment; (5.) Interactions between variables.
- Dependent variable (Y): Screen image brightness.

According to regression standardized residuals of the variables, they are determined whether or not the distribution of the sample is regular. The bell curve is called a completely normal distribution curve. Due to errors of sample, there is a gap between the normal distribution curve histogram and the actual observed-value, as shown in Figure 8. In the experiment, however, there is no discovery for extreme values beyond three standard deviations. The sample values naturally correspond with the normal distribution. Then, the study checks the standardized regression residual error of the variables on the normal P-P plot of regression standardized residuals, as shown in Figure 9.

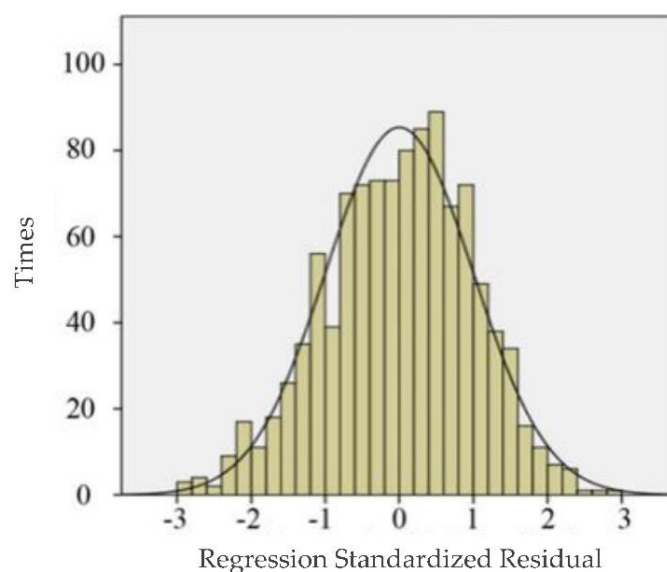


Figure 8. Standardized residuals histogram.

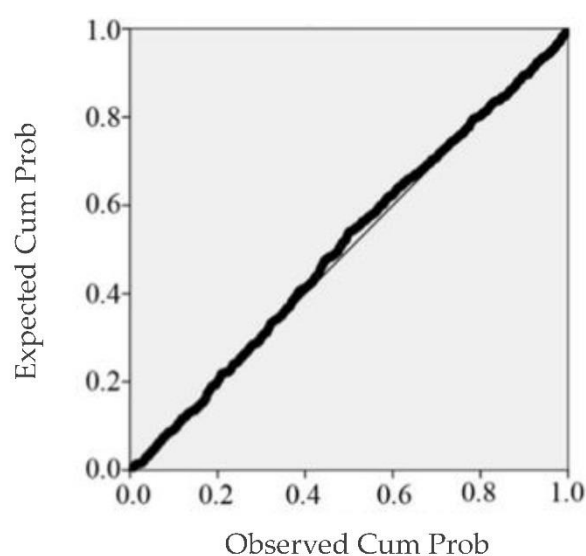


Figure 9. Normal P-P plot of regression standardized residuals.

Table 4 shows the measured coefficients, i.e. correlation coefficient (R); coefficient of determination (R^2); and adjusted coefficient of determination (\underline{R}^2), for adopting in the brightness regression model. The R^2 is utilized to demonstrate the power of the entire pattern in the brightness regression model. Conversely, the measurement tends toward overestimate the phenomenon depending on the sample size; the smaller sample, the more prone model is towards overestimation. Thus, the plurality of researchers uses the adjusted coefficient of determination \underline{R}^2 , that is the error variance and Y variable divided by the degree of freedom.

Table 4. Measured coefficients of the brightness regression model.

Brightness Regression Model	Correlation Coefficient (R)	Coefficient of Determination (R^2)	Adjusted Coefficient of Determination (\underline{R}^2)
	0.995	0.99	0.99

A special form of statistical testing, i.e. analysis of variance (ANOVA) as listed in Table 5, which is widely used to analyze experimental data. Statistical testing is an approach of decision-making depended on the data. If the testing results, i.e. calculated by the null hypothesis, restrict within a certain probability of not being accidental, they are considered to be statistically significant. For instance, the F value is calculated by the data and less than the defined critical significance level. The original hypothesis can be considered invalid. The zero regression coefficient can show that the variable has no influence upon the model.

Table 5. Analysis of variance (ANOVA).

Source	Sum of Squares	Degrees of Freedom	Mean Square	F Value
Return	5382.38	10	538.24	10200.22
Residual	55.61	1054	0.05	
Total	5438.00	1064		

The linear regression model with the statistical coefficients are presented in Table 6. Note that mode of switching between 2D and 3D images has the most important impact on the screen image brightness. In other words, once the screen is exchanged from the 2D to 3D mode, the image brightness on the screen displays a very important decline.

Table 6. Statistical coefficients of linear regression model.

Brightness Regression Variables	Non-Standardized Coefficients		Standardized Coefficients	T
	B	Standard Error	β	
(A) Constant	21.26	0.12		184.89
(1) Screen Size 1	1.91	0.15	0.40	13.18
(2) Screen Size 2	0.71	0.15	0.15	4.89
(B) Shooting Scene Brightness	0.69	0.02	0.15	35.24
(C) 2D or 3D Mode	-5.75	0.12	-1.27	-48.34
(D) Photographic Equipment EV	-1.34	0.01	-1.04	-169.75
(E) Interaction Value of Variables				
(1) and (C)	-0.31	0.03	-0.05	-10.50
(1) and (D)	-0.10	0.01	-0.29	-9.75
(2) and (D)	-0.05	0.01	-0.13	-4.55
(B) and (C)	-0.14	0.03	-0.03	-4.81
(C) and (D)	0.31	0.01	0.99	37.54

The linear equation derived from the method of least squares can get the least-squares line, which is a linear equation that gets the sum of the squared residuals between the data minimal and squared residuals.

The regression equations used in this study are primarily depend on concepts from reference [10], which is used to derive a mathematical model of the basic theory. After transforming the variables of the regression model by Box-Cox to $\lambda = 0.3$, ε represents the normal distribution, as does the R model. The linear regression model parameters are expressed in terms of

$$Y_i = \beta_0 + \beta_i X_i + \varepsilon_i \quad (5)$$

where Y_i is a random variable, X_i is a known fixed constant, ε_i is an unobservable, and $i = 1, \dots, n_x$, whereof i -th test, and Y_i is the reaction value corresponding to X_i . Expressing the primarily variables in the yields of experimental linear model formula

$$Y = 21.259 + 1.905X_1 + 0.707X_2 + 0.69X_B - 5.749X_C - 1.342X_D + \mu \quad (6)$$

where X_1 is the size of screen 1, X_2 is the size of screen 2, X_B denotes the scene brightness of field, X_C denotes the 2D or 3D mode, X_D denotes the EV value of the photographic equipment, and μ is the interaction value of the variables.

This linear regression model is used to analyze the image brightness. The different variables affect brightness of the screen by different degrees, although the interactions on variables affect the screen brightness only minimally. However, changing from 2D to 3D mode has the most important effect on the brightness. Once the screen switched from the 2D to 3D mode, the screen brightness declines obviously.

4.5. Automatic Control of Stereo Camera Stand

The automatic control equipment of the stereo camera stand in experiment consists of seven motors and a circuit board, as shown in Figure 10. According to the parallax information of the stereo imagination which can input and output automatically. The stereo camera stand can be automatically adjusted the distance and angle between the dual cameras by the motors, as shown in Figure 11. When the stereo image is less the feeling of the stereo imagination, the distance of the dual cameras need to extend the distance.



Figure 10. Automatic control equipment of the stereo camera stand in experiment.

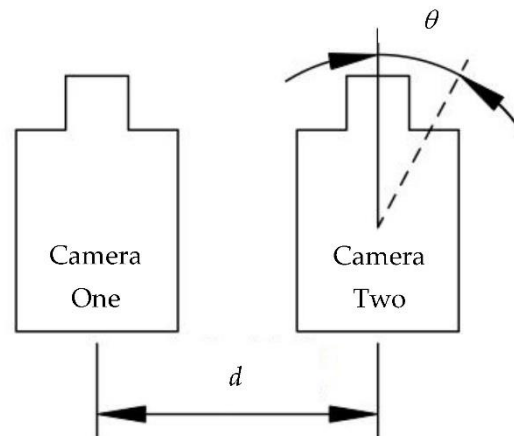


Figure 11. Distance and angle between the dual cameras.

5. Conclusions

An instant preview and analysis system of intelligent stereo cameras (ISCs) is proposed and implemented in this study, and the system has an automatic alignment function before making films and can capture continually stereo video streams with controlled depth of field in real time. The system not only can assist film-makers in reducing filmed time and cost effectively, but also can make sure the quality of stereo films. A novel alignment schedule that considers theoretical accuracy and actual application has been developed. It can utilize in harsh filmed environment, e.g. low light source, indefinite calibrator, and etc. Moreover, a tuning disparity algorithm and control mechanism

with high precision have been developed to achieve correct adjusted results effectively. A prototype with parameter optimization can adopt hardware mechanism to achieve automatic alignment function, and to obtain optimal stereo function by correcting the included gap and angle between dual cameras. Consequently, the system with parameter optimization can enhance the quality of stereo films effectively and can reduce filmed errors and save retouched cost in harsh filmed environment.

Acknowledgments: This work was supported by the Ministry of Science and Technology of Taiwan, R.O.C. under Grant No. MOST 106-2221-E-019-070-. The authors would like to thank the reviewers for their constructive suggestions, which greatly improved this paper.

Author Contributions: Hsing-Cheng Yu conceived and designed the research configuration; Yuan-Zhi Tsai analyzed the data in the experiment; Hsieh-Hung Tsai contributed to implement and setup the experiment. Hsing-Cheng Yu wrote the paper mainly, and all authors read and approved the paper.

Conflicts of Interest: The authors declared that they have no conflicts of interest to this work.

References

1. Bannoa, A.; Ikeuchi, K. Omnidirectional Texturing based on Robust 3D Registration through Euclidean Reconstruction from Two Spherical Images. *Comput. Vis. Image Underst.*, **2010**, *114*, 491–499, 10.1016/j.cviu.2009.12.005.
2. Gao, Z.; Zhang, Y. N.; Xia, Y.; Lin, Z. G.; Fan, Y. Y.; Feng, D. D. Multi-pose 3D Face Recognition based on 2D Sparse Representation. *J. Vis. Commun. Image Represent.*, **2013**, *24*, 117–126, 10.1016/j.jvcir.2012.08.004.
3. Zhang, Y. N.; Guo, Z.; Xia, Y.; Lin, Z. G.; Feng, D. D. 2D Representation of Facial Surfaces, for Multi-pose 3D Face Recognition. *Pattern Recognit. Lett.*, 2012, Vol. 33, pp. 530–536, 10.1016/j.patrec.2011.12.006.
4. Berretti, S.; Bimbo, A. D.; Pala, P. 3D Face Recognition Using Isogeodesic Stripes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2010, Vol. 32, pp. 2162–2177, 10.1109/TPAMI.2010.43.
5. Lenz, R. K.; Tsai, R. Y. Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3-D Machine Vision Metrology, *IEEE Trans. Pattern Anal. Mach. Intell.*, **1988**, *10*, 713–720, 10.1109/34.6781.
6. Lau, M. H.; Armstrong, T. J. The Effect of Viewing Angle on Wrist Posture Estimation from Photographic Images Using Novice Raters. *Appl. Ergon.*, **2011**, *42*, 634–643, 10.1016/j.apergo.2010.08.008.
7. Lowe, B. D. Accuracy and Validity of Observational Estimates of Wrist and Forearm Posture. *Ergon.*, **2004**, *47*, 527–554, 10.1080/00140130310001653057.
8. David, G.; Woodsm V.; Li, G.; Buckle, P. The Development of the Quick Exposure Check (QEC) for Assessing Exposure to Risk Factors for Work-related Musculoskeletal Disorders. *Appl. Ergon.*, **2008**, *39*, 57–69, 10.1016/j.apergo.2007.03.002.
9. Zhang, J.; Shen, S.; Li, L.; Hou, C. A Comparison of Testing Metrics between 3D LCD TV and 3D PDP TV, *Advances on Digital Television and Wireless Multimedia Communications*, **2012**, *331*, 125–132, 10.1007/978-3-642-34595-1_18.
10. Li, P. Box-Cox Transformations: An Overview. Department of Statistics, University of Connecticut, CT, USA, 2005, 3–43.