Auto-positioning mount system for Stereoscopic imaging on variable distances

Zh Yordanov¹, E Fuchs¹, S Löser¹ and J Börcsök¹

¹ Department for Computer Architecture and System Programming, University of Kassel, 34121 Kassel, Germany

E-mail: jivkoj@uni-kassel.de

Abstract. This paper describes a stereo-imaging system for variable distances. The setup includes a mechanical mount system-holder, a control unit for the correct positioning of the cameras, the cameras itself and the object to capture. The goal of this setup was to achieve a true sense of scale and depth of the filmed material on dynamically variable distances to the object. This demands certain requirements which are usually not necessary for many applications of stereoscopic filming. In the last step, the resulting material is prepared for viewing.

1. Introduction

There are a few visual cues which the human brain transforms into volumetric information:

- Perspective – the way in which objects are presented to the eye, based on their spatial attributes.
- Sizes of known objects – the relation of object sizes.
- Detail – closer objects appear in more detail.
- Occlusion – objects, that block another are assumed to be closer
- Lighting – position of light and shadows.
- Relative motion – closer objects seem to move faster, that the objects in background.

These informations enable the extraction of volumetric information even from planar sources. Three more parameters contribute to the volumetric information:

- Binocular disparity – the difference in image location of an object seen by the left and right eyes, resulting from the eyes' horizontal separation.
- Accommodation – the process by which the eye changes optical power to maintain a clear image (focus) on an object as its distance changes.
- Convergence – the simultaneous inward movement of both eyes toward each other, usually in an effort to maintain single binocular vision when viewing an object.

Only the binocular disparity is utilised in stereoscopic imaging.
While binocular disparity is considered the dominant depth cue in most people, if the other cues are presented incorrectly they can have a strong detrimental effect. In order to render a stereo pair one needs to create two images, one for each eye in such a way that when independently viewed they will present an acceptable image to the visual cortex and it will fuse the images and extract the depth information as it does in normal viewing. If stereo pairs are created with a conflict of depth cues then one of a number of things may occur: one cue may become dominant and it may not be the correct/intended one, the depth perception will be exaggerated or reduced, the image will be uncomfortable to watch, the stereo pairs may not fuse at all and the viewer will see two separate images. The degree of the depth factor depends on both the distance of the camera to the projection plane and the separation of the left and right camera. Too large a separation can be hard to resolve and is known as hyper stereo. A good balanced separation of the cameras is 1/30 of the distance to the projection plane. This is generally the maximum separation for comfortable viewing. Another constraint in general practice is to ensure the negative parallax (projection plane behind the object) does not exceed the eye separation. Objects that lie in front of the projection plane will appear to be in front of the computer screen, objects that are behind the projection plane will appear to be "into" the screen. It is generally easier to view stereo pairs of objects that recede into the screen. To achieve this one would place the focal point closer to the camera than the objects of interest [2][3].

2. Method and Algorithms overview

To achieve a correct sense of depth and scale of the stereoscopic content a correct match the viewing geometry with the camera geometry is necessary. As the desired effects depends sensitively on the creation of stereo pairs without conflicting depth cues it is essential to obtain a consistent amount of deviation from the first media to the second. There are mathematical principles that will help this process.

The Bercovitz full formula in equation (1) defines the parameters involved in stereographic calculation [1]. This formula includes both the parallax and depth of field factors to give a maximum permissible stereoscopic effect.

\[ B = P \left( \frac{L \times N}{L - N} \right) \left( \frac{1}{F} - \frac{L + N}{2 \times L \times N} \right) \]  

(1)

\[ L - N = T \]  

(2)

Where \( B = \) Stereo Base (distance between the camera optical axes); \( P = \) Parallax aimed for; \( L = \) Largest distance from the camera lens; \( N = \) Nearest distance from the camera lens; \( F = \) Focal length of the lens; \( T = \) the front to the back thickness of the subject. The equation (2).

The depth factor is based on the amount of deviation (a measure of the difference between the far points and the near of a stereo pair) that looks good in a stereo pair. The most commonly used depth factor used in praxis 1/30. Nevertheless 1/20 and 1/60 the near point distance as a depth factor as also used.

There is a simpler formula in equation (3) given wen the largest distance from the camera is infinity.

\[ B = P \left( \frac{N}{F} - \frac{1}{2} \right) \]  

(3)
In telephoto range, when a subject with reduced depth is recorded the stereo base can be increased. This is only possible for the case that it did not include infinity. The size of this correction can be significant. This kind of imaging is also known as forest photography, where infinity is obscured by the trees. This means stereoscopic infinity can be replaced by the deepest plane in the scene. There are also subjective factors concerning the impression of depth. In some opinions the depth impression should be realistic. Others argue that viewing must be comfortable, but unreal depth is perfectly acceptable, even encouraged. Parallax and stereo depth should be in harmony[4][5].

3. Stereographic system rig
Trial system is constructed by using devices as shown in Figure 1. It includes two stereo cameras, linear drive, balance sheets, step motors, control unit and holder. A standard PC computer is used to control these devices. Stand-alone modus is also provided. Since the recording system is constructed with the components available in the market, initial calibration has to be done.

![Auto-positioning stereoscopic system.](image)

**Figure 1.** Auto-positioning stereoscopic system.

3.1. Mechanics
With the aid of a linear drive, the distance between the cameras can be electrically controlled. They are moving symmetrically and simultaneously (Figure 2). For the volumetric perception it is very important that the angular orientation of the cameras remains precisely controllable. Thus, the slide blocks on the rails have to be as stable as possible while still remaining mobile. A one meter rail is used as a base. The movable mechanism includes a stepper motor and a toothed belt. The mirrored movement of the slide blocks is achieved by attaching the blocks on opposite sides of the belt. Thus it is possible to change the desirable distance between the cameras with a stepper motor without use of additional sensors. To avoid a slipping of the stepper motor, it is necessary to provide a calibration procedure for the position of the cameras. This is realised by triggering the limiter switch and setting this position as zero position. This calibration is done when initially starting the system. A balance plate is provided for accurate alignment to avoid tilting or height differences of the cameras as these
results in a severe disturbance of the volumetric perception. Springs and three screws ensure stable three-point geometry and thus a reliable adjustment. For widening the scope of applications (macro or single object imaging), the possibility to “toe-in” the cameras is also provided. This is achieved using two servo-motors which adjust each other always on opposite angles. The “toe-in” adjustment possibility is still popular in low-cost filming and is included only for compatibility issues. In praxis it should be avoided and the cameras have to be in parallel position.

Figure 2. Linear drive with toothed belt.

3.2. Electrical part
The microprocessor control unit is based on the Atmel AVR architecture. The control board built up of the ATmega644 an 8-bit Atmel® AVR® micro controller. This is combined with a 20MHz Quartz to get up to 20 MIPS. The ATmega644 provides 64kB flash storage for program code, 4kB RAM for temporary data and 2kB eeprom for static data. The In-System Programming (ISP) interface can be used for the first programming via a 10pin connector. The UART0 interface of the micro controller is connected to a MAX232 a RS232-level converter. At the moment, we use this interface for the interaction with the control unit and updating the Firmware via a boot loader. In the future, the Ethernet interface will accomplish both tasks via an ENC28J60 network controller that is connected to the atmega644 by the SPI interface. Four out of an array of up to 32 digital in- and outputs will be used to control the stepper motor end phase. That driver board uses an L298, a dual full-bridge driver IC, in a current limiting circuit. The limit switch is currently contacted to a normal digital input. Regarding the high performance of the atmega644, this poses no problem, but in future the limit switch will be exchanged with an external interrupt port. In the past, we used two potentiometers connected to analog inputs to be able to control the rig without computer. This was removed because it was not necessary after the control via computer was established, but it could be re-integrated to enable faster outdoor tests. (Figure 3 – on the right side).
3.3. Software

The firmware is built in a modular design to achieve easy expandability. It receives strings via the serial connection, interprets them and executes the commands of approach and absolute or relative distance and angle. The firmware is written in C program language and compiled with avr-gcc. The current version consists of the base modules main, stepper, servo, UART and shell. The main module realizes a minimal priority based scheduling. The internal structure is shown in Figure 4 as structogram. The stepper module contains functions to create the half-step sequence and the timing for the stepper motor. Module servo generates two PWM signals (Figure 5) that always provoke contrary angles at the servo motors. It is running in a timer interrupt routine, and thus it is absolutely independent of the scheduling (Figure 6). The functions to receive and transmit chars via the serial-connection are also based on interrupts. These functions take and store the chars in buffers and provide an emergency stop and a reset at any time. But the complex part of the interpretation of the characters and the strings is subject to the scheduling. This is stored in the shell module and it is the most intricate part of the firmware at the moment. The shell provides two modes. The first one is a key mode in which single characters are interpreted as simple commands. The second mode is a command mode that stores the characters in a buffer until a character has the "return" value. Then the string parses for known commands and absolute or relative values. Relative values are initialized by a plus or minus as prefix of the value.

![Internal software algorithm](image-url)

**Figure 4.** Internal software algorithm.
4. Conclusion
We presented an auto-positioning mount concept for stereoscopic imaging, based on binocular stereo vision. The system allows a high range of parameters and works fast and reproducible. There is a wide scope of potential applications. For example, if we add a zoom control we could be able to zoom in continuously while keeping the stereoscopic effect alive which is not possible with conventional devices. This continuous zoom enables a virtual movement toward an object.

References
[1] Bercovitz, J. 1998 Image-side perspective and stereoscopy J. Stereoscopic Displays and Virtual Reality Systems
[2] Dyer. C.R. 2001 Volumetric scene reconstruction from multiple views J. Foundations of Image Analysis.
[3] Groth, S 1983 Three dimensional and stereoscopic graphics for scientific data display and analysis J. IEEE Computer graphics and Applications. 3,8, 1983, pp 31-34
[4] Hodges L.F and, McAllister D.F. 1985 Stereo and alternating pair techniques for display of computer generated images J. IEEE Computer Graphics and Applications. 5, 9, 1985, pp 38-45
[5] M Goesele, N.Snively, B.Curless, H.Hoppe and S.M. Seitz 2007 Multi-view stereo for community photo collections J. ICCV
[6] M. Okutomi, T. Kanade 1993 A multiple-baseline stereo. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 15, No. 4, April, 1993, pp. 353-363.