Analysis and design of stereoscopic television systems

AM Ariyaeeinia

University of Hertfordshire,
College Lane, Hatfield, Hertfordshire AL10 9AB, UK

Tel: (+44) 1707-284348
Fax: (+44) 1707-284199
E-mail: A.M.Ariyaeeinia@herts.ac.uk

Analysis and design of stereoscopic television systems

A. M. Ariyaeeinia
University of Hertfordshire, Hatfield, Hertfordshire, AL10 9AB, UK

Abstract. The performance of stereoscopic television systems and its dependence on various parameters are investigated. It is shown that the extent of the region of space which can be displayed in three dimensions without causing undue eyestrain to the observer, and also the minimum detectable depth interval, are related to and may therefore be controlled by the system geometrical and optical parameters. Distortion in three-dimensional images due to the difference between the depth magnification and the magnification of the xy-plane is also analysed. Finally, it is demonstrated that the expressions defining the relationship between the performance of a 3D television system and its various parameters form a useful basis for the design of a system of this type for a given application.

1. Introduction

Three-dimensional television systems have been the subject of extensive research for many years [1,4-6,8,10-12]. In a 3D television system, a pair of spatially rotated and spaced apart cameras is employed to provide the observer with views of the same scene from two different angles. Certain important advantages of these systems over two-dimensional viewing have made them the superior choice in a variety of applications. Examples of these include remote inspection, co-ordinate measurement, telemanipulation, and guidance of mobile robots [7,9,13-15]. For any given application, the correct operation of the system depends on how well its characteristics satisfy the requirements of the task. Since such requirements may vary considerably for different applications, it would seem necessary that the design of a stereoscopic television system should be formulated in accordance with the specific aims involved. It is believed that this may be achieved successfully if the relationship between the performance features of a three-dimensional television system and its various parameters are directly incorporated into the design.

This paper concentrates on certain important characteristics of a three-dimensional television system which together, or otherwise, specify its performance in a wide range of applications. These are the extent of the comfortable stereo-region, the minimum detectable depth interval, and the image magnification. For the purpose of this investigation a three-dimensional television system has been designed and implemented which is of a sequentially-switched type. However, since the treatment is general, the results are applicable to all stereoscopic television realisations.

2. Extent of the stereo-region

The overlapped fields of view of a pair of TV cameras, which are spatially rotated and spaced apart to view the same scene from different angles, form a bounded volume in 3D object space known as the stereo-region. Any object point in this region, provided it is not occluded, has two spatially separated images on the monitor screen. The existing parallaxes between corresponding image points observed separately through the left and right eyes give rise to a three dimensional perception of the picture.

Figure 1 illustrates the geometry of a stereoscopic imaging system when the TV cameras are rotated such that their central axes converge at a point C. The extent of the stereoscopic region can be evaluated as the difference between the ranges, from the camera base-line, of the furthest and nearest points in this region. With the aid of Figure 1 it can be easily seen that the range, from the camera base line, of any point (A) on the system central axis is give by

$$Z_A = \frac{B(P + 2f\tan\alpha)}{4f - 2P\tan\alpha},$$
(1)

where B is the camera separation, f is the focal length of the stereo-camera lenses, α is the rotation angle of the cameras, and P is the horizontal parallax on the image plane ($P = l - [-r]$, or $P = 2r$ since the point is on the central axis of the system).

For the furthest and nearest points (G and H) in the stereo-region, the parallaxes are equal to K and $-K$ respectively, where K is the width of the stereo-camera targets. The extent of the stereo-region is therefore given by:

$$Z_G - Z_H = \frac{B(K + 2f\tan\alpha)}{4f - 2K\tan\alpha} - \frac{B(-K + 2f\tan\alpha)}{4f + 2K\tan\alpha}$$
(2)

or

$$Z_G - Z_H = \frac{2BKf(1 + \tan^2\alpha)}{4f^2 - K^2\tan^2\alpha}.$$
(3)

It should be pointed out that for $\alpha = \tan^{-1}\frac{2f}{K}$, the furthest point in the stereo-region, Z_G, and therefore the depth of the stereo-region extend to infinity. The reason is that for this value of α, the lines defining the outer boundaries of the two fields of view (Figure 1) become perpendicular to the camera baseline and therefore parallel with each other. This can be seen by noting that θ in Figure 1 is given by $\tan^{-1}\frac{K}{2f}$. As a result, for the above value of α, the angle between each of the outer boundary lines and the camera baseline ($\alpha + \theta$ in Figure 1) becomes equal to $\frac{\pi}{2}$. For values of α between $\tan^{-1}\frac{2f}{K}$ and $\frac{\pi}{2}$, these outer boundary lines in fact diverge from one another, and the extent of the stereo-region remains infinite. In such cases, it is clear that (3) can no longer be considered valid for calculating the stereo-region extent. The particular case in which α is equal to $\frac{\pi}{2}$ is discussed in section 6.1.
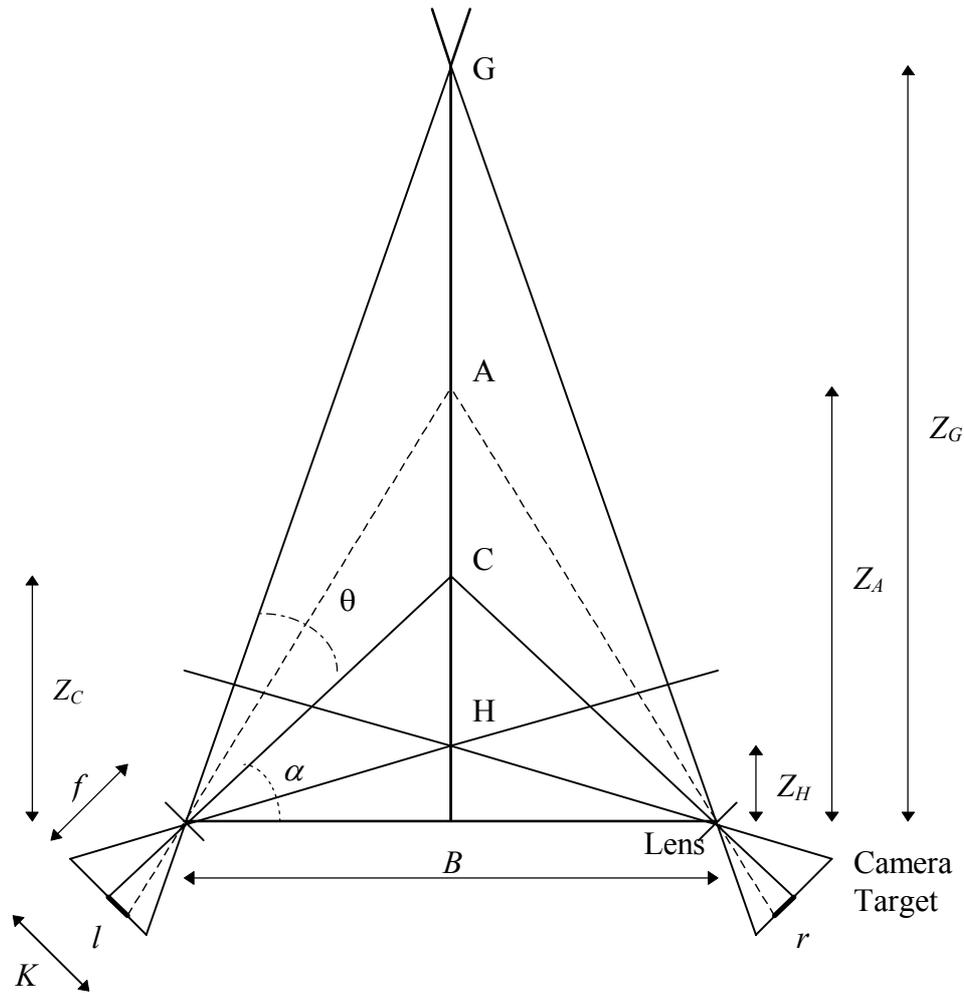
Figure 1. Geometry of the stereoscopic imaging.

3. Comfortable stereo-region

It has been demonstrated [1] that the region of space which can be displayed in three dimensions without causing undue eyestrain to the observer does not cover the total overlap region of the fields of view of the stereo-pair of TV cameras. This is due to the fact that in order for an observer to fuse the left and right images and therefore see the picture in three dimensions, the horizontal parallax on the monitor screen must not exceed a certain maximum value. This limitation is partly a result of the finite depth of field of the eye and the displacement of the apparent position of the reconstructed 3D image from the monitor screen. When the observer's eyes are focused on the monitor screen, the depth of field of his eyes covers only a finite distance in front and also beyond the monitor screen plane. The parallax between the corresponding images on the monitor screen, however, causes the central axes of the observer's eyes to converge at a point either behind or in front of the monitor screen plane. This convergence point is where the observer sees the 3D image. When the apparent location of the 3D image is around either the far or near edge of the depth of field of the eye, the observer may only fuse the two corresponding images with considerable

effort which will then result in strain to his eyes. For reconstructed images located well outside the depth of field of the observer's eyes, the fusion is completely lost and the observer sees two images instead of one in three dimensions.

Another reason for the loss of fusion is the wide parallax between the corresponding image points on the observer's left and right retinas. When an observer looks at a point in a 3D image, a pair of corresponding retinal images are formed at identical locations in the observer's two eyes. The images of points other than the watched points, however, are not always reconstructed on the corresponding locations on the two retinas. This effect is known as binocular retinal disparity and is the most important cue for depth perception [5]. It has been noted that [10,12] if the retinal disparity exceeds a certain limit then the left and right images cannot be fused and double vision takes the place of a stereoscopic perception.

Based on the above discussion, in order to determine the extent of the comfortable stereo-region, the horizontal parallax in (3) should be replaced with the maximum permissible parallax on the image plane. Denoting this latter parallax by P_m, the extent of the comfortable stereo region, ΔZ, is obtained as

$$\Delta Z = \frac{2BP_m f(1 + \tan^2 \alpha)}{4f^2 - P_m^2 \tan^2 \alpha} . \tag{4}$$

This equation indicates that for $\alpha = \tan^{-1} \frac{2f}{P_m}$, the comfortable stereo-region extends to infinity. The reason behind this is that, in this case, the lines defining the left and right boundaries of the comfortable stereo-region become parallel with each other. It should also be emphasised that (4) is not valid for values of α greater than $\tan^{-1} \frac{2f}{P_m}$. However, as indicated below, P_m is much smaller than the focal length. As a result, the upper limit on the useful value of α is usually very close to $\frac{\pi}{2}$.

The maximum permissible parallax on the image plane, P_m, can be calculated using the maximum tolerable parallax on the monitor screen, P'_m. i.e.

$$P_m = -\frac{P'_m}{M}, \tag{5}$$

where M is the magnification factor from the image plane to the monitor screen. It has been demonstrated experimentally that P'_m is a user dependent parameter, directly related to the viewing distance [1,4]:

$$P'_m = WD, \tag{6}$$

where D is the viewing distance and W is the observer-related coefficient with a typical value of 0.018 for an average viewer. Using this value of W and a viewing distance of 1 m, Figures 2 to 4 illustrate the way the extent of the comfortable stereo-region varies with each of the parameters f, B and α.
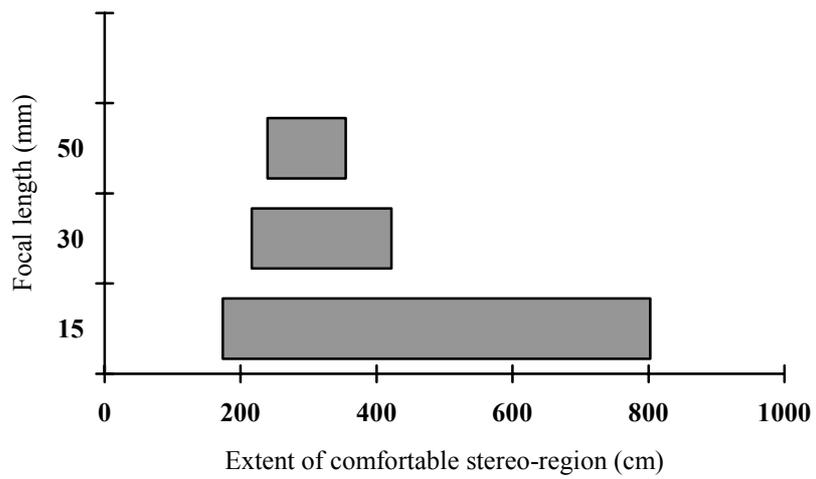
Figure 2. Effect of the focal length of the stereo-camera lenses on the depth and distance from the camera base line, of the comfortable stereo-region.

$B = 20$ cm, $\alpha = 88°$, $M = 26.66$
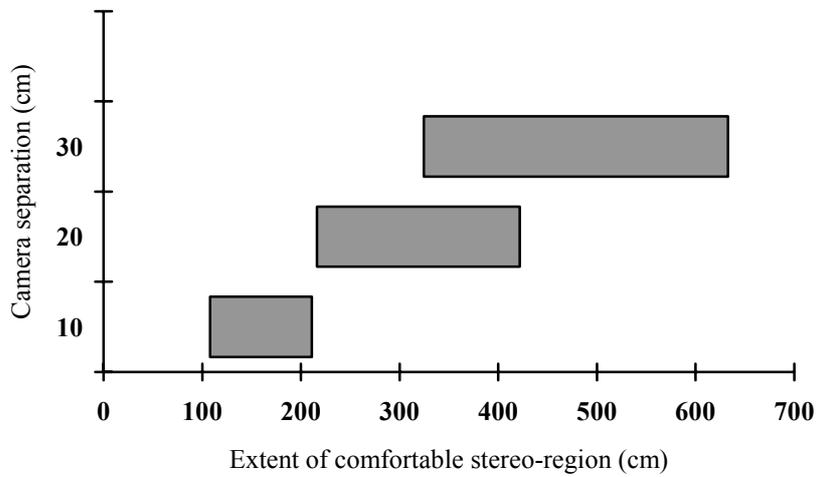

Figure 3. Effect of the camera separation on the depth and distance from the camera base line, of the comfortable stereo-region.

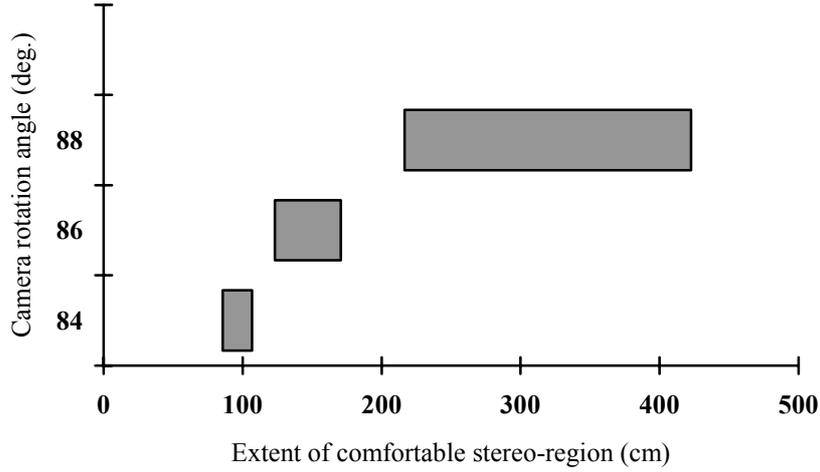$f = 30$ mm, $\alpha = 88°$, $M = 26.66$

Figure 4. Effect of the camera rotation angle on the depth and distance from the camera base line, of the comfortable stereo-region. $f = 30$ mm, $B = 20$ cm, $M = 26.66$

It is evident from the above figures that whereas any change in the focal length only varies the extent of the comfortable stereo-region, variation in either the camera separation or rotation angle will both change the depth of, and shift the location of, this region. To show the reason behind this, Figure 1 can be used to write an expression for the range from the camera base line to the system convergence point (point C):

$$Z_C = \frac{B \tan \alpha}{2}.$$

(7)

The above equation clearly indicates that the location of point C and therefore that of the comfortable stereo-region are affected only by the camera separation and rotation angle.

4. Minimum detectable depth interval

In many practical applications, e.g. remote inspection and telemanipulation, the accuracy of the operation is highly dependent on the minimum detectable depth-interval, or depth resolution. The fact that depth perception is due to the existence of disparity between the left and right images suggests that the minimum detectable depth interval is directly related to the minimum detectable parallax on the image plane which itself is determined by the resolution of the stereo-camera targets.

To obtain an expression for the depth resolution, (1) can be used to write

$$\delta Z = \frac{4Z^2 f (1 + \tan^2 \alpha)}{B(P + 2f \tan \alpha)^2} \delta P.$$

(8)

Taking δP as the minimum detectable parallax on the image plane, δZ represents the depth resolution around a range Z in object space. It has been indicated that, in practice, the camera rotation angle must be very close to $\frac{\pi}{2}$ in order to ensure that the vertical parallax in the stereo-image is negligible [3]. In this case the depth resolution is given by

$$\delta Z = \frac{Z^2}{Bf} \delta P,\qquad(9)$$

since $2f \tan\alpha \gg P$, and $\tan^2 \alpha \gg 1$.

Evidently the minimum detectable depth interval is not constant along the stereo-region and increases with the range squared. The above equation further implies that δZ can be improved by increasing the camera separation and the focal length of the stereo-camera lenses. Figure 5 gives an example of the variation of δZ with range.
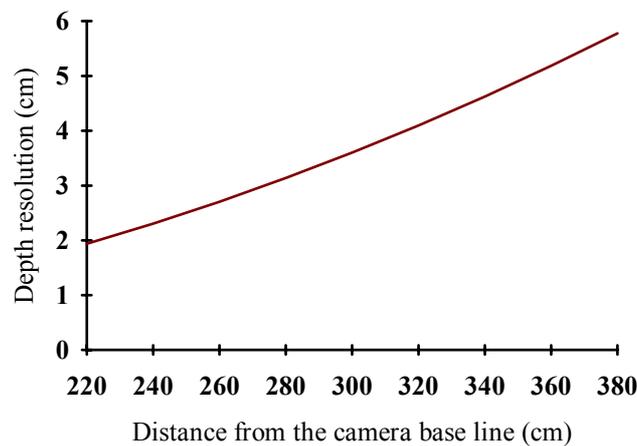

Figure 5. Variation of the depth resolution with range.
$f = 30$ mm, $B = 20$ cm, $\delta P = 24 \ \mu$m

5. Image magnification

In a stereoscopic television system magnification has two distinct aspects, i.e. the depth magnification and the magnification of the xy-plane. Since the values of these are not necessarily identical [3], the display may suffer from some undesired level of distortion. This is particularly unacceptable in applications involving visual inspection, where the realism of the 3D image is of prime importance.

The magnification of the xy-plane on the monitor screen (m_{xy}) and that on the image plane ($\tilde{m}_{xy}$) are related as

$$m_{xy} = M\tilde{m}_{xy},\qquad(10)$$

where M is again the magnification factor from the image plane to the monitor screen. Since $\tilde{m}_{xy}$ is equal to the ratio of the camera focal length to the object range, (10) can be expressed as [2,3]

$$m_{xy} = \frac{Mf}{Z}.$$ (11)

The magnification of depth on the other hand may be stated as

$$m_z = \frac{\delta d}{\delta Z},$$ (12)

where δZ is a small depth variation in object space and δd is the corresponding depth change in the display. Using (9) it is possible to find the relationship between δZ and the corresponding increment in the horizontal parallax on the monitor screen ($\delta P'$):

$$\delta Z = -\frac{Z^2 \delta P'}{BfM},$$ (13)

since $\delta P' = -M\delta P$.

Figure 6 illustrates the reconstruction of an image point away from the monitor screen. The distance between the apparent location of this point and the observer is

$$d = \frac{ED}{P' + E},$$ (14)

where E is the separation of the observer's eyes, and P' is the horizontal parallax on the monitor screen. Using the above equation it is possible to write

$$\delta d = -\frac{ED}{(P'+E)^2}\delta P'.$$ (15)

However, since the maximum tolerable parallax on the monitor screen is much less than the distance between the observer's left and right eyes, the above equation reduces to

$$\delta d = -\frac{D\delta P'}{E}.$$ (16)

Taking $\delta P'$ as a small variation of the horizontal parallax on the monitor screen, δd represents the corresponding depth change in the apparent location of the 3D image point. Substituting (13) and (16) in (12), the following expression for the depth magnification is obtained:
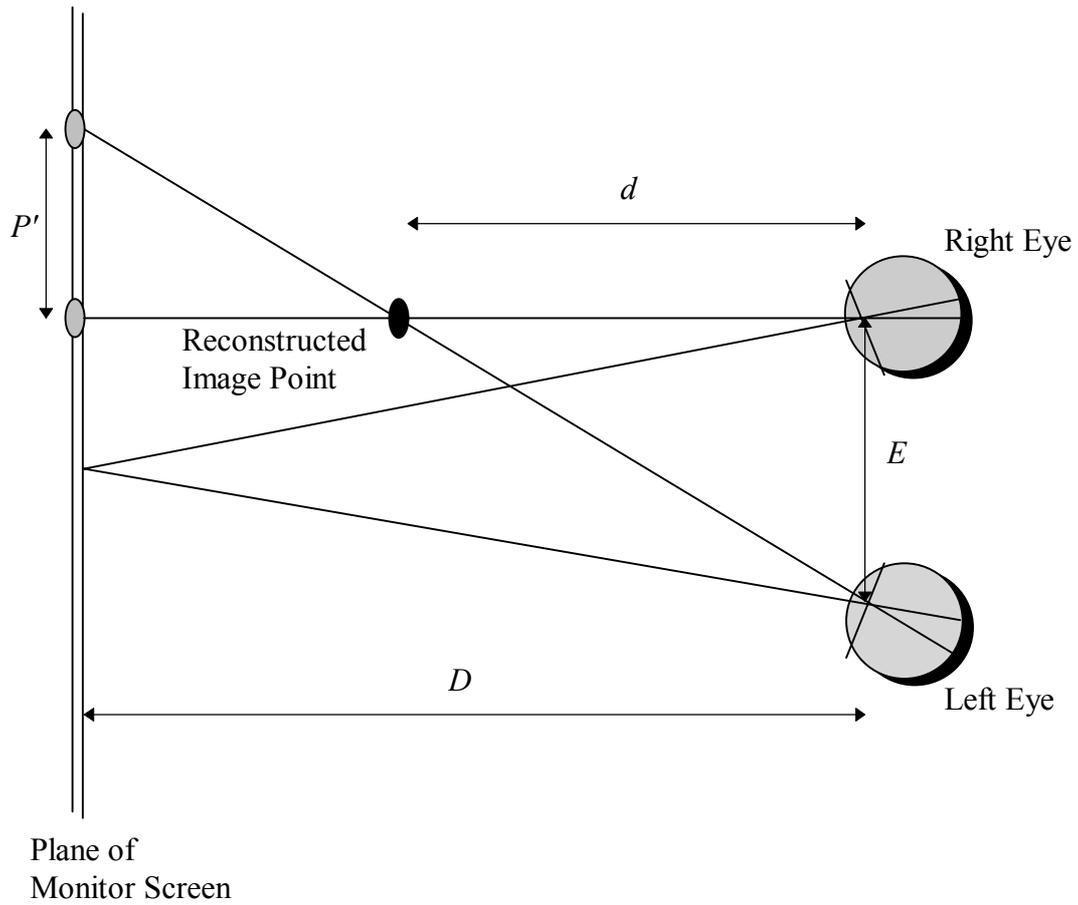
$$m_z = \frac{DBfM}{EZ^2}.$$ (17)

Figure 6. Apparent location of the reconstructed image point.

The ratio of the depth magnification to the magnification of the *xy*-plane is therefore given by

$$r_m = \frac{m_z}{m_{xy}} = \frac{DB}{EZ}.$$
(18)

This equation indicates that for a fixed camera separation and viewing distance, the magnification ratio changes with object range, and is equal to unity only for $Z = DB/E$. It is therefore necessary that the parameters B and D should be chosen such that, for the volume of interest within the comfortable stereo-region, the deviation from unity of r_m is not large enough to result in an intolerable degree of distortion in the display.

6. Proposed design procedure

The problem of designing a stereoscopic television system can be defined as one of determining the system parameters such that the required three-dimensional performance is obtained. Some of these parameters, however, can not always be considered as variables. For instance, the minimum resolvable parallax on the image

plane, δP, and the magnification factor from the image plane to the monitor screen, M, are determined by the type of video equipment available. Furthermore, the parameters related to the human eye (i.e. E and W) should also be regarded as fixed. For any given design case, the values of all such parameters must be obtained a priori and then used in the evaluation of the remaining system parameters.

For the purpose of this study, the application of 3D TV systems in remote inspection is considered. It is assumed that the object under inspection is to be located on the convergence point in the stereoscopic region, and also that the parameters M, δP, E, W and D are fixed at some appropriate values. Finally, the system performance must meet the requirements of the task specified in terms of the minimum extent of the comfortable stereo-region, the depth resolution and the magnification ratio. The design procedure in this case can have the following form.

I) Using (18), the values of Z_C and B are determined. This may be done by replacing r_m and Z in that equation first with 1 and Z_C, and then with $(1+\delta r_m)$ and $(Z_C - \hat{Z})$ respectively, where δr_m is the maximum permissible deviation from unity of the magnification ratio, and $\hat{Z}$ is the estimated extent of the volume of interest in front of the system convergence point:

$$\frac{DB}{EZ_C} = 1 \tag{19}$$

and

$$\frac{DB}{E(Z_C - \hat{Z})} = 1 + \delta r_m. \tag{20}$$

Using these equations the lower limit on Z_C is first determined. Evidently, the smaller δr_m is and the larger $\hat{Z}$ is, the larger the lower limit on Z_C becomes. By deciding a value for Z_C accordingly, it can be assured that the deviation from unity of the magnification ratio at $Z_C - \hat{Z}$ is less than the maximum permissible value. By inserting the decided value for Z_C in (19), B can be calculated such that the magnification ratio at the system convergence point is equal to unity.

II) The value of f is found by replacing Z in (9) with Z_F, where Z_F is the distance from the camera base line of the object furthest point (or the furthest point of interest within the comfortable stereo-region):

$$f = \frac{Z_F^2}{B\delta Z}\delta P. \tag{21}$$

III) By inserting the calculated values of B and Z_C in (7), the rotation angle of the stereo-camera, α, can be computed as

$$\alpha = \tan^{-1}\left(\frac{2Z_C}{B}\right). \tag{22}$$

IV) The maximum permissible parallax in the image plane, P_m, can be determined by combining (5) and (6). This may then be used in the evaluation of the depth of the comfortable stereoscopic region and also in evaluating the extent of this region both beyond and in front of the system convergence point.

It should be pointed out that if it is required to increase the extent of the comfortable stereoscopic region, a convenient method is to reduce the focal length of the stereo-camera lenses. This assures that the magnification ratio requirement will still be satisfied and that, the location of the comfortable stereoscopic region will not be affected.

6.1. Parallel TV cameras

The discussions presented above are for the general form of the stereoscopic structure in which the two cameras are rotated such that their central axes converge at a point, C. By setting the stereo-camera rotation angle (α) to $\frac{\pi}{2}$, however, a specific form of the stereoscopic structure is obtained in which the axes of the cameras are in parallel with each other and the two image planes (camera targets) are collinear. In this case the system convergence point is at infinity. This means that the 3D image is always formed in front of the plane of the monitor screen and never behind it. An advantage of the parallel positioning of the TV cameras is that it results in a zero vertical parallax in the stereo-image [3].

With such a geometry of TV cameras, (1) can be rewritten for any point in the stereo-region as

$$Z_A = \frac{Bf}{P} \tag{23}$$

and the stereo-region extends from Z_H to infinity, where Z_H in this case is given by

$$Z_H = \frac{Bf}{K}. \tag{24}$$

Similarly, the comfortable stereo-region extends from the nearest point in this region to infinity. The range from the camera baseline of the nearest point in the comfortable stereo-region is given by

$$Z_{nearest} = \frac{Bf}{P_m}. \tag{25}$$

It should also be pointed out that for parallel TV cameras, in fact, (9) becomes exact and the expression given in (18) for the magnification ratio still applies.

For the purpose of design, Z_C in (19) and (20) should be replaced with Z_L which denotes the range from the camera baseline of a point in the stereo-region considered as the centre of inspection volume. The magnification ratio associated with this point should be set to unity. Having calculated the values of B and Z_L, the focal length, f, and the maximum permissible parallax on the image plane, P_m, can be obtained as before. P_m may then be used to evaluate the range from the camera baseline of the nearest point in the comfortable stereo-region.

7. Conclusions

An investigation into the performance features of three-dimensional television systems, and their relationships with various parameters, has been presented. The study has shown that, through the use of the system geometrical and optical parameters, it is possible to adjust the depth resolution and the extent of the comfortable stereo-region in order to satisfy the requirements of a given task. Display distortion due to the difference between the depth magnification and the magnification of the *xy*-plane has also been discussed. This class of distortion detracts from the realism of the display and hence has undesired effects in applications involving visual inspection. It has been demonstrated that the difference in the two types of magnification varies with the object range and can be controlled using the camera separation and viewing distance. Based on the equations specifying the above relationships, a procedure for the design of a stereoscopic television system for use in one of its main application areas, i.e. remote inspection, has been developed. It is believed that the approach presented can also be used as a model in the design of such a system for a variety of other applications.

8. References

[1]   A.M. Ariyaeeinia, "The Design and Performance of Stereoscopic Television Systems", *Proc. 4th SPIE Conf. on Mobile Robots*, Philadelphia, USA, November 1989, Vol. 1195, pp. 362-370.

[2]   A. M. Ariyaeeinia, "Calibration of an Active Stereoscopic Imaging System", *Proc. 6th SPIE Conf. on Optics, Illumination and Image Sensing for Machine Vision*, Boston, USA, November 1991, Vol. 1614, pp. 146-154.

[3]   A. M. Ariyaeeinia, "Distortions in Stereoscopic Displays", *Proc. 3rd SPIE Conf. on Stereoscopic Displays and Application*, San Jose, USA, February 1992, Vol. 1669, pp. 2-9.

[4]   A. M. Ariyaeeinia, "Analysis of 3-D TV Systems", *Proc. SPIE Conf. on Display Systems*, Munich, Germany, June 1993, Vol. 1988, pp. 11-17.

[5]   R. Buffet, "The Visual Perception of Depth Using Electro-Optical Display Systems", *Displays*, 1980, pp. 39-45.

[6]   W. N. Charman, "Three-Dimensional TV Systems", *Applied Opt.*, 1981, Vol. 20, No. 1, pp. 14-16.

[7]   A. Dumberk, "Depth of Vision: 3D TV", *IEE Review*, March 1993, pp. 61-64.

[8]   M. G. Maxwell, "Three Dimensional Color Television", *SMPT J.*, August 1966, pp. 821-825.

[9]   A. Meregalli, "Telepresence Application of Real Task Efficiency Improvement in a wide and Hostile Environment", *Proc. SPIE Conf. on Telemanipulator and Telepresence Technologies*, Boston, USA, November 1994, Vol. 2351, pp. 148-159.

[10] N. Ogle, "Stereoscopic Depth Perception and Exposure Delay between Images to the two Eyes", *J. Opt. Soc. Am.*, Vol. 53 No. 11, 1963, pp. 1296-1304.

[11] Okoshi, *Three Dimensional Imaging Techniques*, Academic Press Inc., London, 1976.

[12] A. Roses and A.S. Khalafalla, "Stereoscopic Viewing with PLZT Ceramics", *Ferroelectrics*, Vol. 10, 1976, pp. 47-51.

[13] H. Takemura, A. Tomono and Y. Kobayashi, "Study of Pseudo-Workspace Using a Stereoscopic Display", *Proc. 4th SPIE Conf. on Mobile Robots*, Philadelphia, USA, 1989, Vol. 1195, pp. 371-377.

[14] N. Tetsutani and F. Kishino, "3-Dimensional Display Method Without Special Glasses for Virtual Space Teleconferencing System", *Proc. SPIE Conf on Display Systems*, Munich, Germany, June 1993, Vol. 1988, pp. 18-25.

[15] D. J. Wenzel, S. B. Seida and V. R. Sturdivant, "Telerobot Control Using Enhanced Stereo Viewing", *Proc. SPIE Conf. on Telemanipulator and Telepresence Technologies*, Boston, USA, November 1994, Vol. 2351, pp. 225-232.