Resampling radially captured images for perspectively correct stereoscopic display

N. A. Dodgson

University of Cambridge Computer Laboratory,
Gates Building, J. J. Thomson Avenue, Cambridge, UK CB3 OFD

This paper appears in *Proc SPIE* **3295** "SPIE Symposium on Stereoscopic Displays and Applications IX", 24th-30th Jan 1998, San Jose, California, pp. 100–110

ABSTRACT

When rendering or capturing stereoscopic images, two arrangements of the cameras are possible: radial ("toed-in") and parallel. In the radial case all of the cameras' axes pass through a common point; in the parallel case these axes are parallel to one another. The radial configuration causes distortions in the viewed stereoscopic image, manifest as vertical misalignments between parts of the images seen by the viewer's two eyes. The parallel case does not suffer from this distortion, and is thus considered to be the more correct method of capturing stereoscopic imagery. The radial case is, however, simpler to implement than the parallel: standard cameras or renderers can be used with no modification. In the parallel case special lens arrangements or modified rendering software is required. If a pinhole camera is assumed it should be readily apparent that the same light rays pass through the pinhole in the same directions whether the camera is aligned radially to or parallel to the other cameras. The difference lies in how these light rays are sampled to produce an image. In the case of a non-pinhole (real) camera, objects in focus should behave as for the pinhole case, while objects out of focus may behave slightly differently. The geometry of both radial and parallel cases is described and it is shown how a geometrical transform of an image produced in one case can be used to generate the image which would have been produced in the other case. This geometric transform is achieved by a resampling operation and various resampling algorithms are discussed. The resampling process can result in a degradation in the quality of the image. An indication of the type and severity of this degradation is given.

Keywords: stereoscopic, autostereoscopic, resampling, image processing, camera, geometry

1. INTRODUCTION

When capturing or rendering stereoscopic images two arrangements of the cameras tend to be used: the radial ("toed-in") and the parallel.[1–3] These are illustrated in Figure 1.

In the radial case all of the cameras' axes pass through a common point: the point of convergence. In the parallel case these axes are parallel to one another. The radial case is simpler to implement, but produces stereoscopic misalignments in the viewed imagery which make it difficult for the viewer to stereoscopically fuse the images[4] (Figure 2). The parallel case does not produce these misalignments (Figure 3). It can thus be considered the correct way to capture stereoscopic images. The parallel case is, however, generally more difficult to implement: it requires either a wider-angle lens than the radial case, accompanied by clipping of the resulting image (Figure 4) or an off-axis setting of the camera's lens (Figure 5), requiring special mounting of the lenses.

Images produced by radially arranged cameras can be processed to provide the images which would have been produced by cameras arranged in parallel. This provides the advantages of the parallel case without the implementation difficulties described above. The method, described below, is applicable to any number of cameras from two upwards.

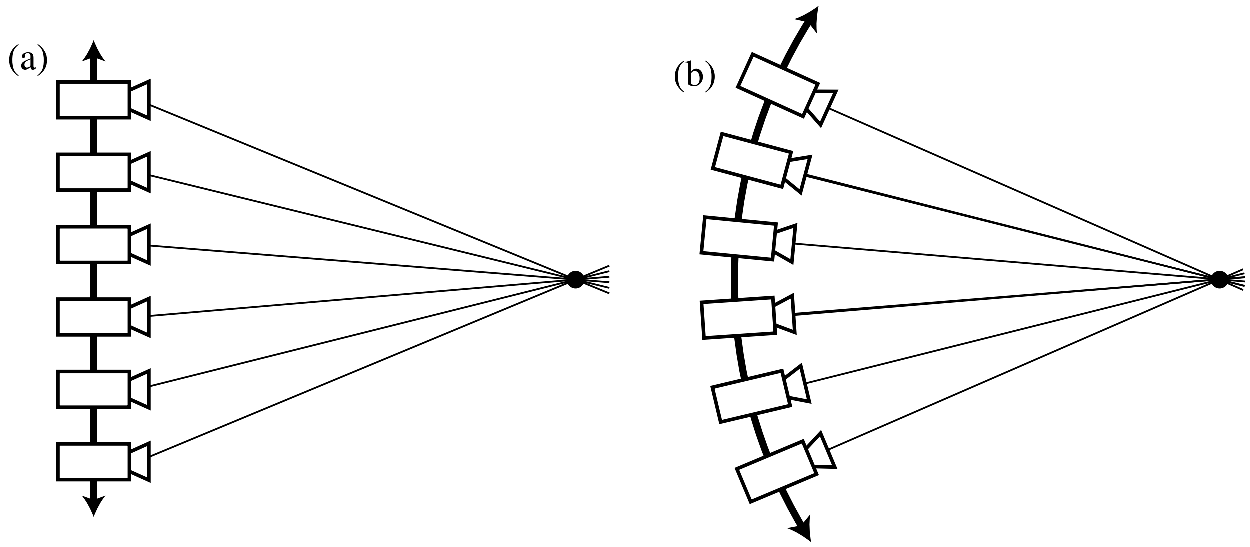

Figure 1. (a) the parallel configuration; (b) the radial configuration.

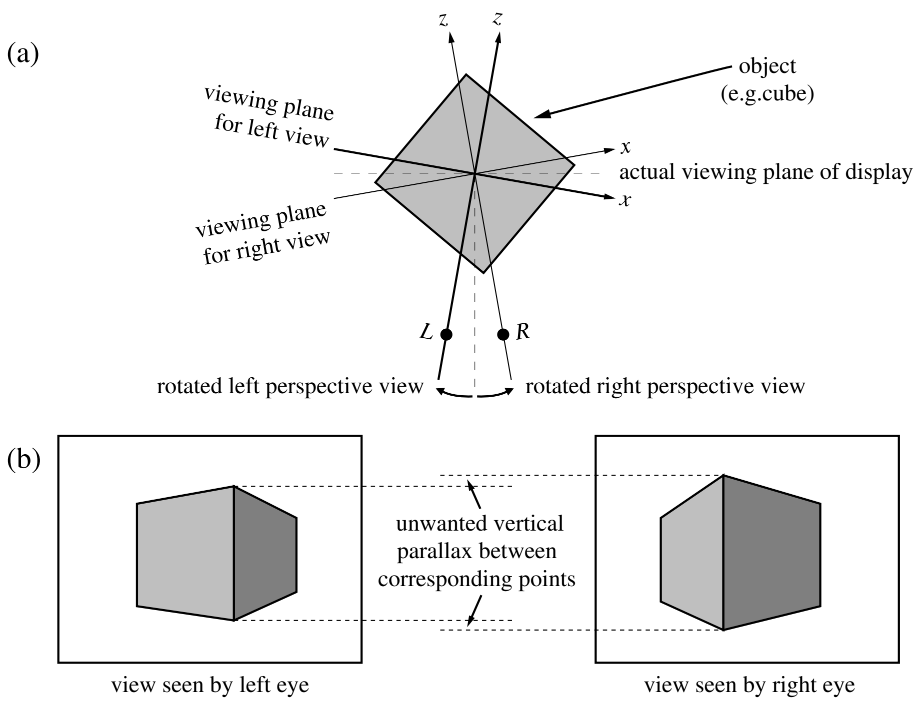

Figure 2. Radial configuration: (a) the geometry of rotation of perspective views for stereo; (b) distortions in the stereo image due to rotation of the perspective views. These distortions are caused by displaying both rotated views on the display's unrotated viewing plane.

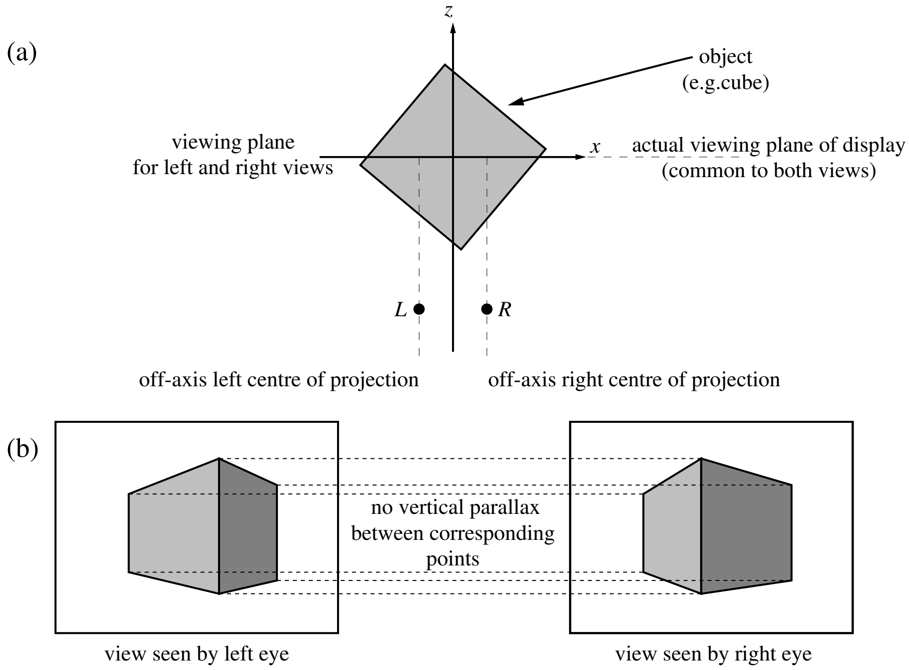

Figure 3. Parallel configuration: (a) the geometry of parallel capture of perspective views for stereo; (b) correct stereo image resulting from parallel capture of perspective views.

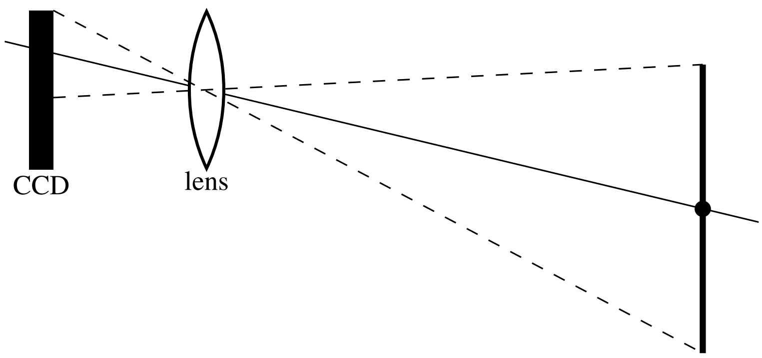

Figure 4. When capturing images in the parallel configuration with conventional cameras, only part of the CCD images the desired scene. The resulting image needs to be clipped to ensure that only the appropriate part is used.

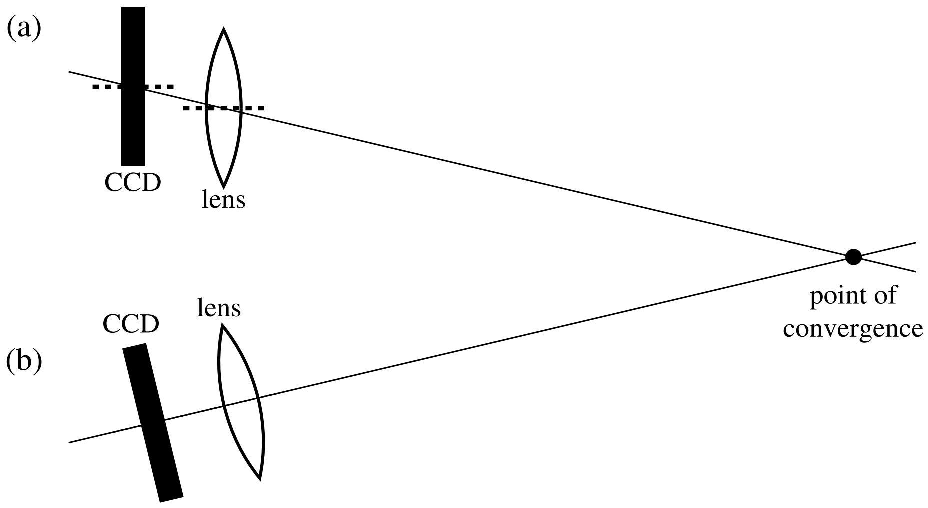

Figure 5. (a) an off-axis setting of the camera lens can be used instead of the inefficient arrangement shown in Figure 4. Compare this with (b), the radial arrangement, where a standard camera can be used with no modification and with no loss of part of the CCD.

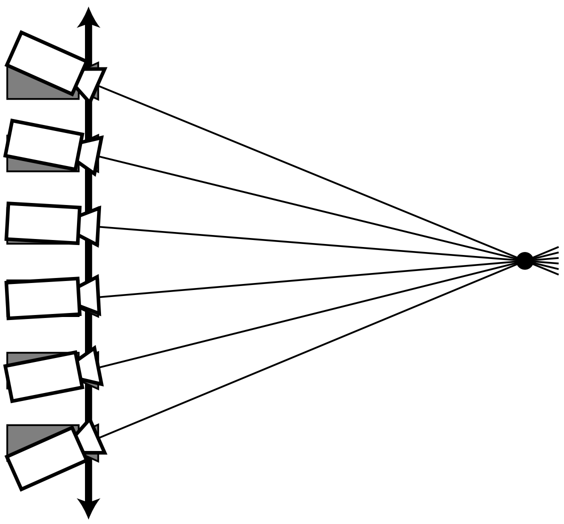

Figure 6. Cameras are arranged as for the parallel case (Figure 1(a)). Each is rotated about the appropriate principal point of its lens system so that their axes all pass through the point of convergence.

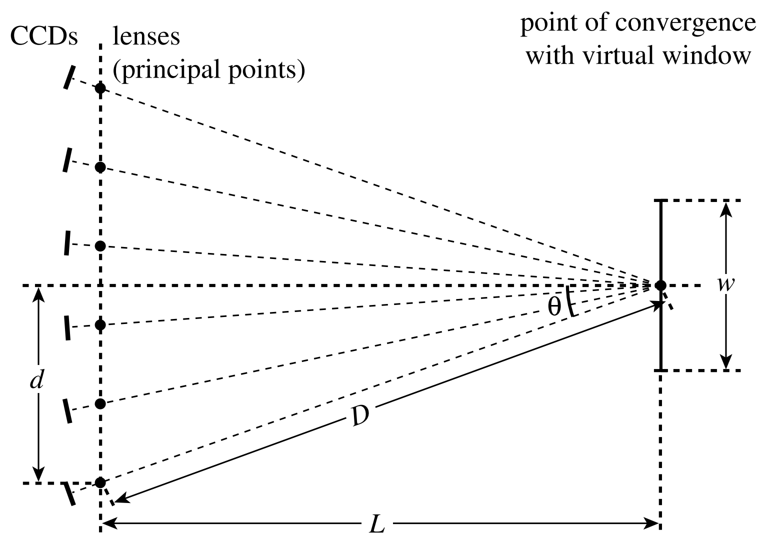

Figure 7. The overall system. The lenses' principal points are arranged in a straight line, a distance L from the point of convergence. They are equispaced along this line. In the parallel case each CCD would take an image of the virtual window (width w). In this radial approximation each CCD will take an image of the whole virtual window with some margin around the edge owing to the fact that the virtual window maps to a trapezoid on the rotated CCD. When considering a single camera we need to know d which, with L, gives us also D and θ.

2. METHOD

Cameras are positioned such that the appropriate principal point[5] of each camera's lens system is in the position it would be in the parallel case (Figure 6). The cameras are, however, aligned so that their axes all pass through the point of convergence, in a similar fashion to the radial case. Note that the optical centres are in a straight line, as in the parallel case, and not on a circular arc. The geometry of the system is shown in Figure 7. The images which would have been produced in the parallel case can be generated from the images actually produced by the cameras using an image resampling process, which can be implemented in either hardware or software or some combination of the two.

It can be shown mathematically that the rectangular image which would have been captured in the parallel case is present in the radially captured image in a perspectively distorted trapezoid form. By inverting this distortion, corrected images can be obtained.

If we assume pinhole cameras then it should be readily apparent that the same light rays pass through the pinhole in the same directions regardless of the camera's alignment. Figure 8 shows the geometry for a single camera. To convert the radially captured image to its parallel equivalent we need to find the projection from one image's coordinate system to the other's. We essentially project the image from the radial plane onto the parallel plane. Each of the two planes has its own local 2D coordinate system with origin at the centre of the image. These origins can be thought of as the optical image of the point of convergence of the cameras.

One potential difficulty of this geometry is the fact that the resampling operation appears to need to know the size of the CCD and its distance from the appropriate principal point of the camera's lens or from the pinhole in the ideal case. However, the geometry to the left of the pinhole is congruent to that on the right. We can, therefore, perform all of our resampling operations as if the images were formed on planes passing through the point of convergence, as shown in Figure 9. A final transformation will be required to obtain pixel coordinates rather than physical coordinates.

Other author information:
E-mail: nad@cl.cam.ac.uk
HTTP: http://www.cl.cam.ac.uk/users/nad/

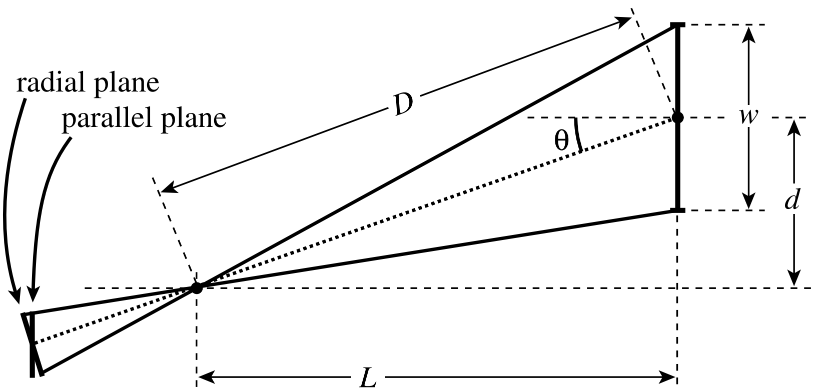

Figure 8. The geometry for a single camera showing the important parameters. Assuming a pinhole camera, the same light rays hit the radial plane as hit the parallel plane; they are merely sampled differently. Note that the geometry to the left of the pinhole is congruent to that on the right.

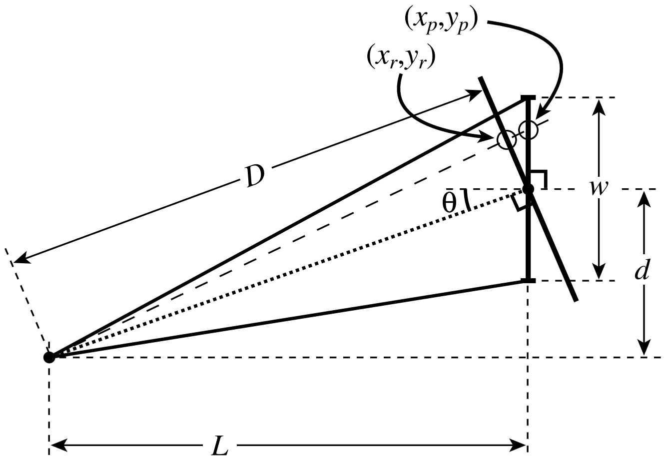

Figure 9. The geometry used in this analysis. Both radial and parallel planes pass through the point of convergence. A ray from the pinhole intersects the radial plane at (x_r, y_r) and the parallel plane at (x_p, y_p). These are coordinate systems local to each plane. The origins of both these coordinate systems are at the point of convergence, with the y-axes perpendicular to the plane in which lie the point of convergence and the cameras' principal points.

A careful analysis of this geometry leads to a simple relationship between the radial coordinates, (x_r, y_r), and the parallel coordinates, (x_p, y_p).

$$x_p = x_r \frac{1}{\cos\theta} \frac{1}{1 - \frac{x_r}{D}\tan\theta} \tag{1}$$

$$y_p = y_r \frac{1}{1 - \frac{x_r}{D}\tan\theta} \tag{2}$$

$$x_r = x_p \cos\theta \frac{1}{1 + \frac{x_p}{D}\sin\theta} \tag{3}$$

$$y_r = y_p \frac{1}{1 + \frac{x_p}{D}\sin\theta} \tag{4}$$

This leaves the problem of determining D and θ from the more easily measured parameters d and L. D and L are always positive but d and θ may be either positive or negative. The above equations imply that θ is positive in the example shown in Figure 9. This means that d is positive below the point of convergence and negative above it, where "above" and "below" refer to the orientation shown in Figures 7, 8 and 9. We can therefore write:

$$\begin{aligned} \theta &= \tan^{-1}(d/L) \\ D &= \sqrt{d^2 + L^2} \end{aligned}$$

and also:

$$\begin{aligned} \sin\theta &= d/D \\ \cos\theta &= L/D \\ \tan\theta &= d/L \end{aligned}$$

thus allowing us to transform equations (1) through (4) into a different form if required.

Given this geometric mapping between the parallel and radial cases, it is possible to resample an image from one case to produce a good approximation to the other.

3. RESAMPLING

3.1. Fundamentals

An image is essentially a rectangular array of evenly spaced samples. This contrasts sharply with the conventional view of an image as an array of abutting square pixels. You can reconcile these two models of an image by thinking of the samples as being taken at the centres of the square pixels. The "abutting square pixels" representation is, in fact, only one of many possible physical realisations of the underlying sample values. Display engineers, for example, will know that a CRT and an LCD produce quite different physical realisations of the pixels.

Resampling involves taking one of these arrays of samples and generating a new array whose samples are taken in different locations. In practice this is achieved by taking each sample point in the new image, geometrically transforming it to find its location in the old image, and then using the old sample values nearest to that location to construct a value for the new sample (Figure 10). In essence a continuous function is reconstructed from the discrete old sample values and new samples are taken from this function at the appropriate locations.[6–8] Resampling thus consists of a reconstruction function, which generates a continuous function from discrete sample points, and a sampling operation, which generates new sample points from this continuous function.

3.2. Reconstruction Functions

The simplest of the reconstruction functions is nearest-neighbour: every point in the continuous function takes on the value of the nearest sample point. This, though fast, produces unwanted artifacts in the resulting image, as shown in Figure 11. A slower method, with a more pleasing result, can be generated if we take the weighted mean of the four sample points nearest to the location of the new sample (Figures 12 and 13). This is bi-linear interpolation.

Bi-linear interpolation, while better than the nearest neighbour, tends to blur the resulting image. Even better results can be obtained by considering the 3×3 nearest (bi-quadratic) or 4×4 nearest (bi-cubic) sample points.[6,9–12]

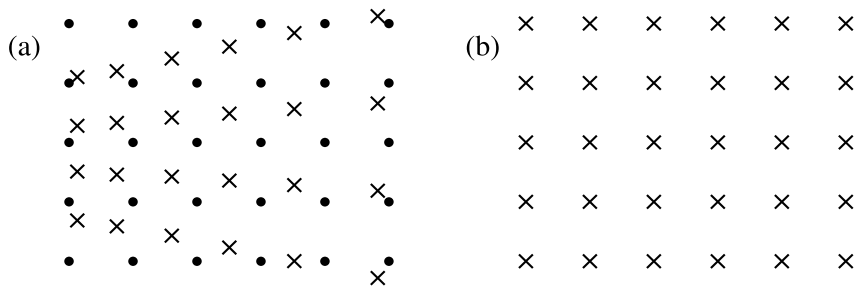

Figure 10. The circles represent sample points in (a), the original image, while the crosses represent sample points in (b), the resampled image. Each sample point in (b) is transformed to its equivalent location in (a) and the original sample points in the immediate vicinity are used to generate the resampled value.

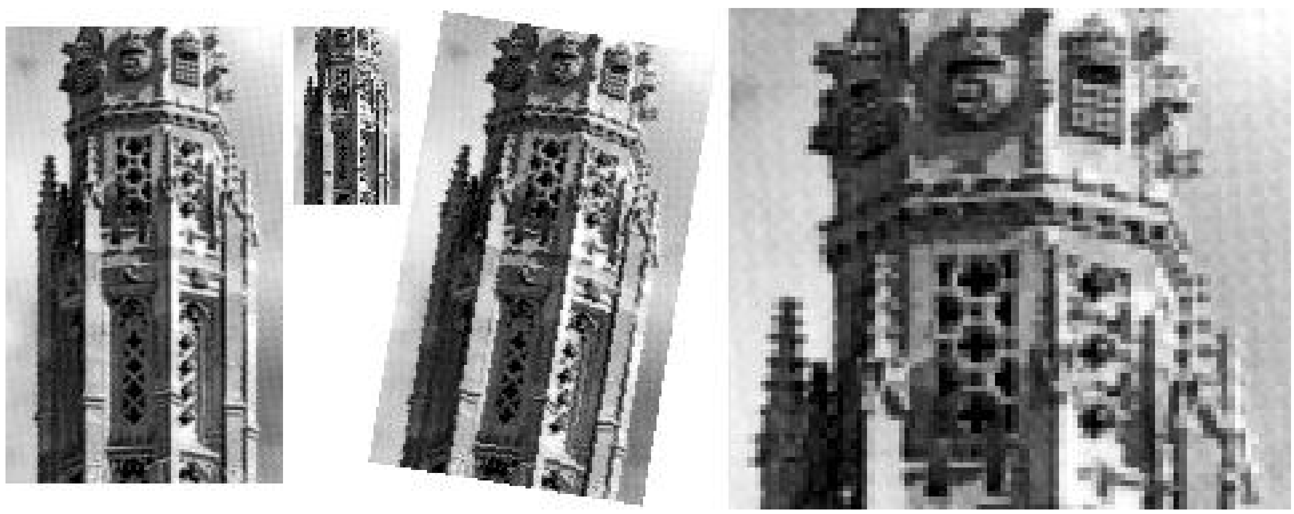

Figure 11. Examples of the artifacts which occur with nearest-neighbour reconstruction. At left is the original. At right, an enlarged version, clearly showing the block structure. In the middle are a reduced version, showing the problems of undersampling, and a rotated version, showing artifacts along the vertical edges of the tower.

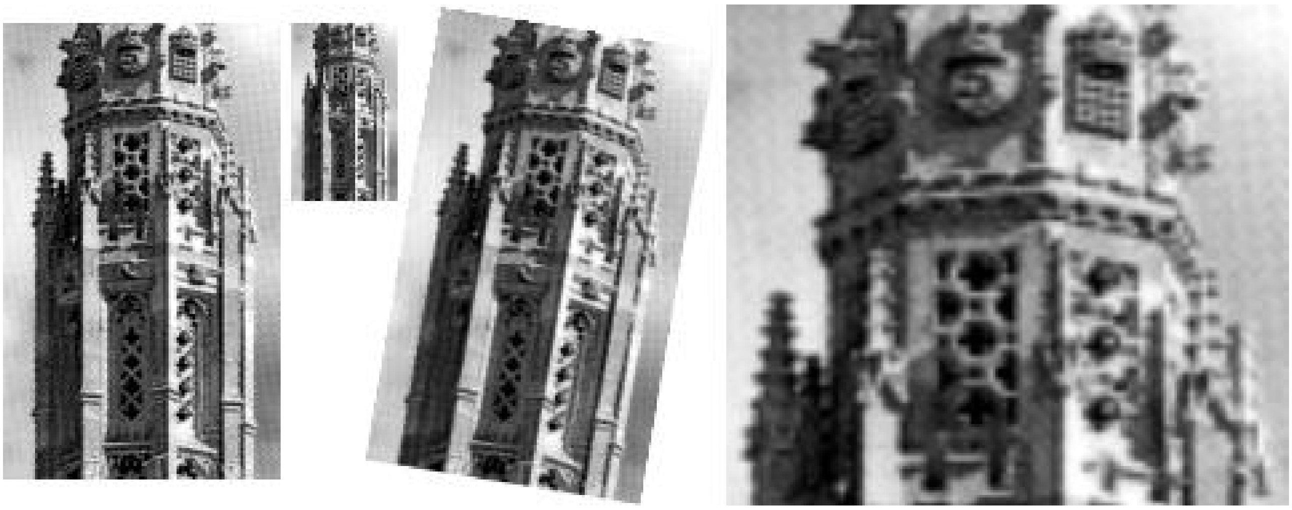

Figure 12. The same examples under linear reconstruction. The artifacts are significantly reduced, at the expense of some blurriness.

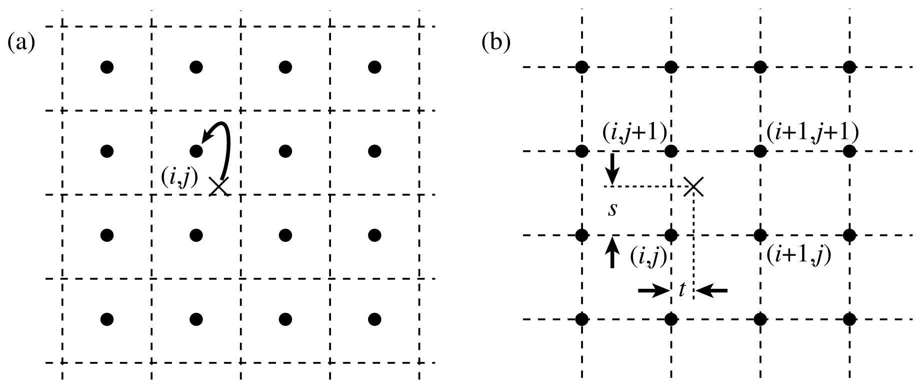

Figure 13. (a) nearest-neighbour reconstruction takes the nearest sample value to the transformed sample point while, (b), linear reconstruction takes a weighted mean of the four nearest sample points $F'(x, y) = (1 - t)(1 - s)F_{i,j} + (1 - t)sF_{i,j+1} + t(1 - s)F_{i+1,j} + tsF_{i+1,j+1}$.

This can obviously continue to any number of points, but beyond cubic little quality is added to the resulting image. Indeed it has recently been shown that the cubic is only marginally better than the quadratic while requiring 60% extra computation time.[13]

There are other reconstruction functions besides these local polynomials: truncated Gaussians, variants on the sinc function, global polynomials, and more.[8] These, however, require considerably more computation than the local polynomials and, for our purposes, give no advantage over these simpler functions.

3.3. More Complex Sampling

There are situations in which it is advantageous to use more complex sampling than simply a single point sample for each pixel. Digital cameras and scanners do not perform point sampling; each sample is an average over some small area. It may, therefore, be thought that area-averaged sampling could be performed in the resampling operation. This is, unfortunately, very difficult to implement. A tractable approximation to area sampling is to take multiple point samples throughout the area and to average these. In computer graphics this general technique is known as super-sampling and has a number of subtle variants.[14] The precise nature of the resampling operation determines whether or not super-sampling is required in preference to single point sampling.

3.4. Trade-offs

The use of single point sampling with nearest-neighbour reconstruction is the fastest possible resampling method. It will produce an image which approximates to the ideal result, but which is likely to contain undesirable artifacts (Figure 11). The artifacts take different forms depending on the local scale differences between the original and resampled images. In enlarged areas the resampled image contains an array of regions of constant colour. In reduced areas the resampled image under-samples the information in the original: leading to unwanted aliasing artifacts. Even in areas where the scales of the original and resampled images are roughly the same this resampling method can still produce unsightly effects, especially near sharp edges in the image.

The solution to these problems is to use a better resampling method which, while unable to produce the perfect result, will nevertheless ameliorate the artifacts. It has been shown that, for regions of image reduction, it is the quality of the sampling operation which dominates the quality of the overall resampling operation. On the other hand, for same-size or image magnification operations, it is the quality of the reconstruction function which dominates the quality of overall resampling.[8]

One obvious solution would be always to use a good super-sampler (e.g. 4×4 samples per pixel) with a good reconstruction function (e.g. bi-quadratic or bi-cubic). This, however, leads to unnecessary computation in the resampling operation: the super-sampler is not necessary when magnifying and the good reconstruction function is not necessary when reducing. A reasonable compromise between quality and efficiency can be effected whereby super-sampling and nearest-neighbour reconstruction are used in areas where there is an overall local reduction in scale of, say, two or more; while single point sampling and bi-quadratic reconstruction are used in areas of roughly the same scale or magnification.

Magnification of the original is not desirable because it involves stretching the known information over a wider area. When we magnify real-world objects we see more detail. When we magnify a digital image we see the original sample data at a larger scale and we get no extra detail. This destroys the vital illusion that the digital image is a valid representation of the real world. On the other hand, significant reduction of image requires the capture of many more pixels than are really necessary for a good result.

Ideally, therefore, the original image should be of such a scale that the greatest reduction factor is no more than two and the least no less than one. Here the benefit of super-sampling is debatable and single point sampling with bi-linear or bi-quadratic interpolation is likely to be acceptable.

4. NON-PINHOLE CAMERAS

The analysis has considered the ideal case of a perfect pinhole camera. A real camera does not behave ideally because it has a lens of non-zero size. Parts of the image that are in focus will form as in the ideal pinhole case. These will resample correctly between the two configurations.

Parts of the image which are out of focus will be blurred slightly differently by the two configurations (Figure 14) and, we suspect, will not be equivalent under resampling. In practice this small difference is immaterial, as objects which are out of focus are generally of little importance to the task in hand.

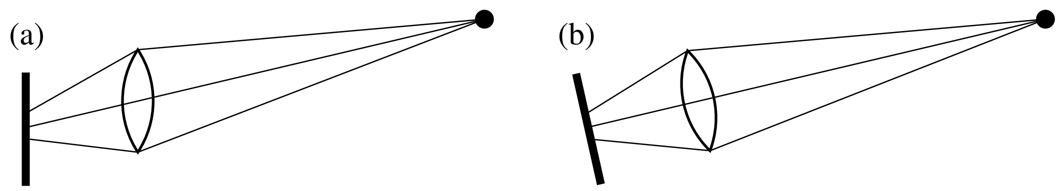

Figure 14. The image of an out of focus point formed in (a) the parallel case and (b) the radial case.

5. CONCLUSION

It is possible to configure standard cameras in a radial arrangement to produce the images which would have been generated by cameras arranged in parallel. Image resampling is used to generate one set of images from the other. This allows us to capture perspectively correct stereoscopic images without the need for special cameras.

ACKNOWLEDGMENTS

Thanks to Jenni Cartwright for typing the original manuscript and to Dr Oliver Castle for Figures 2 and 3.[4] Part of this work was funded by Autostereo Systems Limited of Cambridge, UK.

REFERENCES

1. N. A. Dodgson and J. R. Moore, "Time-multiplexed autostereoscopic camera system", *Proc. SPIE*, **3012**, pp.72–83, 1997.
2. B. Choquet, C. Guerin, A. Dumbreck, and P. M. Scheiwiller, "Development of a European stereoscopic camera for television imaging", *Proc. SPIE*, **1901**, pp.37–45, 1993.
3. L. F. Hodges and D. F. McAllister, "Computing stereoscopic views", *Stereo computer graphics and other true 3D technologies*, D. F. McAllister (ed), Princeton University Press, pp.71–89, 1993.
4. O. M. Castle, *Synthetic Image Generation for a Multiple-View Autostereo Display*, Technical Report No. 382, Computer Laboratory, University of Cambridge, Pembroke Street, Cambridge CB2 3QG, UK, Oct, 1995.
5. W. T. Welford, *Optics* (second edition), Oxford University Press, 1981.
6. D. P. Mitchell and A. N. Netravali, "Reconstruction filters in computer graphics", *Comp. Graphics*, **22**(4), pp.221–288, 1988.
7. G. Wolberg, *Digital Image Warping*, IEEE Computer Society Press Monograph, IEEE Computer Society Press, 10662 Los Vaqueros Circle, P.O.Box 3013, Los Alamitos, CA 90720–1264, USA; 1990.
8. N. A. Dodgson, *Image Resampling*, Technical Report No. 261, Computer Laboratory, University of Cambridge, Pembroke Street, Cambridge CB2 3QG, UK, Aug, 1992.
9. H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering", *IEEE Trans. Acoust., Speech, Sig. Proc.*, **26**(6), pp.508–517, 1978.
10. R. G. Keys, "Cubic convolution interpolation for digital image processing", *IEEE Trans. Acoust., Speech, Sig. Proc.*, **29**(6), pp.1153–1160, 1981.
11. J. A. Parker, R. V. Kenyon and D. E. Troxel, "Comparison of interpolation methods for image resampling", *IEEE Trans. Med. Imag.*, **2**(1), pp.31–39, 1983.
12. E. Maeland, "On the comparison of the interpolation methods", *IEEE Trans. Med. Imag.*, **7**(3), pp.213–217, 1988.
13. N. A. Dodgson, "Quadratic interpolation in image resampling", *IEEE Trans. Image Proc.*, **6**(9), pp.1322–1326, 1997.
14. J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics: Principles and Practice*, Addison-Wesley, 1990.