PROCEEDINGS REPRINT

SPIE—The International Society for Optical Engineering

Reprinted from

Human Vision, Visual Processing, and Digital Display III

10–13 February 1992
San Jose, California

Volume 1666

Color spatial acuity control of a screen subdivision image synthesis algorithm

Gary W. Meyer
Aihua Liu

Department of Computer and Information Science
University of Oregon
Eugene, OR 97403

ABSTRACT

The limited color spatial acuity of the human visual system is exploited to develop a more efficient algorithm for realistic image synthesis. A screen subdivision ray tracer is modified to control the amount of chromatic and achromatic detail present at the edges in an environment. An opponents color space (previously used to select wavelengths for synthetic image generation) is used to define the chromatic and achromatic channels present in the image. Computational savings achieved by the algorithm are discussed. A perceptual evaluation shows that image quality is not seriously degraded by the use of the technique.

1. INTRODUCTION

The human visual system is a remarkable imaging device. Many of its properties, such as its wide dynamic range and its ability to interpret three dimensional forms, have not been successfully duplicated in today's photographic cameras and machine vision algorithms. However, there are also limitations to the imaging capabilities of the visual perceptual system. Certain of these restrictions have been explained as filters that help us select the important information from our visual field and allow us to better function in our environment. The reasons for other deficiencies have not been made clear, but some appear to be the result of shortcomings inherent in most biological optical systems.

The limitations of the visual system can be taken advantage of in realistic image synthesis. If a piece of information lies below the threshold of visibility then it makes no sense to spend time and effort computing it. This fact has long been understood in the fields of image processing and image transmission, and many algorithms have been developed in these areas that improve image quality and lower data transmission rates by taking the characteristics of the visual perceptual system into account. Synthetic image generation techniques in computer graphics can also minimize computational expense and decrease storage requirements by capitalizing on what is known about how the visual system works.

This paper presents a realistic image synthesis technique that takes advantage of the limited color spatial acuity of the human visual system. It is based upon an opponents processing model of color vision that is consistent with previous work that we have done in minimizing the number of wavelengths used for color calculations[1] and with proposals that have been made for storing and transmitting color images. The paper begins by reviewing what is known about why color vision has limited spatial resolution. Next the modifications that were made to an existing screen subdivision ray tracer to implement the technique are described. Finally the computational advantages of the algorithm are discussed and a psychophysical analysis of the approach is presented.

2. CHARACTERISTICS OF HUMAN COLOR SPATIAL VISION

The human visual system exhibits different spatial acuities for information that comes from different portions of the spectrum or different locations in color space. The complete reason for this variation in color spatial acuity is not completely understood, but several of the contributing factors have been identified. In this section we discuss what is known about the artifacts that are introduced at various stages of the visual pathway, from chromatic aberration in the optical system of the eye, through the limited spatial sampling of the receptor system, and finally to the different spatial frequency responses of the opponent color pathways.

2.1. Chromatic aberration

Because the index of refraction for a material is a function of wavelength, each wavelength of light that enters the human eye is focussed to a different depth relative to the retina. The shorter wavelengths of light are focussed in front of the retina and the longer wavelengths are focussed in back. This is called axial chromatic aberration, and it limits the resolution to which the eye can resolve each wavelength in the spectrum.

Bedford and Wyszecki[2] give the focal length (expressed in diopters) of the lens that is necessary to correct the eye at each wavelength. As can be seen in Figure 1, their results show that the correction required varies from about -2.0 diopters at 400 nm to about 0.5 diopters at 700 nm. Given the amount of defocussing that occurs it is possible to compute the corresponding modulation transfer function that is appropriate at each wavelength.[3,4] Figure 2 shows how spatial frequencies are attenuated as the amount of defocussing increases.

2.2. Receptor spacing

After passing through the optical elements of the eye, light that reaches the retina (under photopic viewing conditions) is detected by the cone receptor system. Each cone in the retina has one of three different spectral sensitivities. A spectral energy distribution is thus reduced to three pieces of information by the time it leaves the retina. If two different spectral energy distributions produce signals with the same tristimulus composition, then the observer will say that they are the same color. This is the basis of all color reproduction work in computer graphics.

The CIE XYZ color matching functions that are used in most color reproduction calculations are related by a lineai transform to the fundamental spectral sensitivity functions of the receptor system. Given the dichromatic confusior points for the three major types of color defective vision this transformation from CIE XYZ space to SML space car be specified. Using the confusion points suggested by Estevez[6] this transformation becomes[7]

$$\begin{bmatrix} S \\ M \\ L \end{bmatrix} = \begin{bmatrix} 0.0000 & 0.0000 & 0.5609 \\ -0.4227 & 1.1723 & 0.0911 \\ 0.1150 & 0.9364 & -0.0203 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (1$$

Figure 3 shows the short (S), medium (M), and long (L) wavelength spectral sensitivity functions that result fror applying this transformation to the CIE XYZ matching functions.

There is increasing evidence regarding the numbers and the spacing of the SML receptors on the retina. In recer work it was assumed that the separation between cone centers in the fovea centralis is 0.6 minutes of arc.[8] This mear that there would be $60/0.6 = 100$ receptors per degree of vision or that the maximum spatial frequency that could b sampled without aliasing is 50 cycles/degree. These numbers are consistent with what is known about the modulatio transfer function of the human visual system.[9] Given that the spacing of the S receptors is approximately one in every 1 minutes of arc,[10] there must be $60/10 = 6$ S receptors and therefore $100 - 6 = 94$ M and L receptors per degree of visio It is also known that there are approximately twice as many L receptors as M receptors.[8] Therefore, the approxima number of each type of receptor in the retina is 62 L, 32 M, and 6 S per degree of vision.

2.3. Spatial acuity of opponent channels

It is now well established that color assumes an opponent representation immediately following the receptor stage of tl human color vision system. Evidence of this fact comes from psychophysical,[11,12] neurophysiological,[13] and theoretical studies that have been done. In previous work[1] we have used the opponent fundamentals to select the wavelengths which to perform realistic image synthesis.

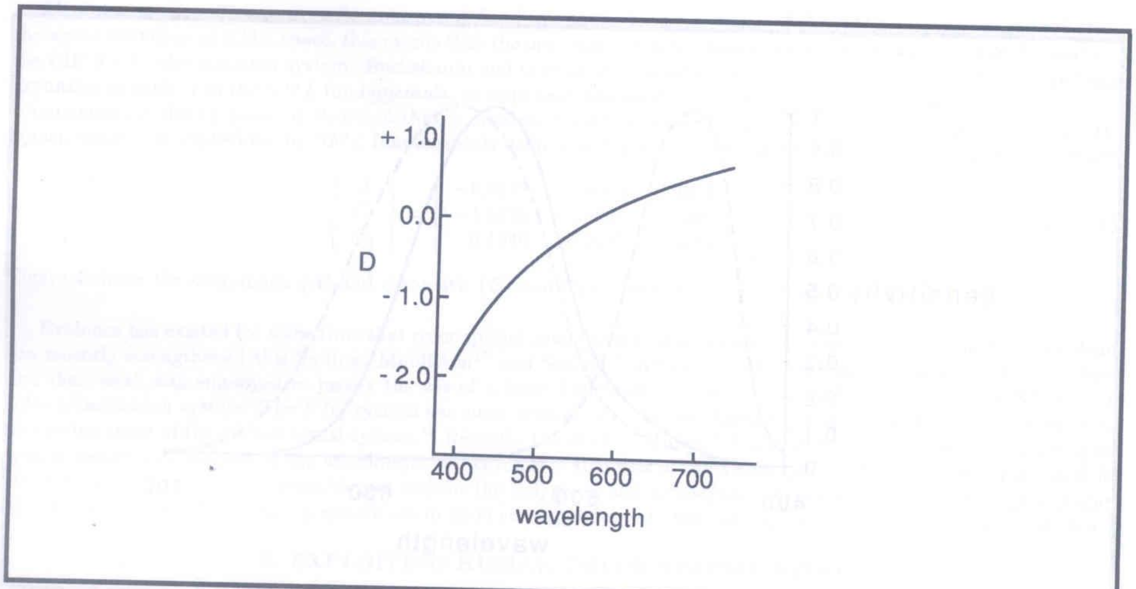

Figure 1: Change in focal length (expressed in diopters) necessary at each wavelength to overcome chromatic aberration in the eye (from Bedford and Wyszecki[2]). The reference wavelength is 578 nm.

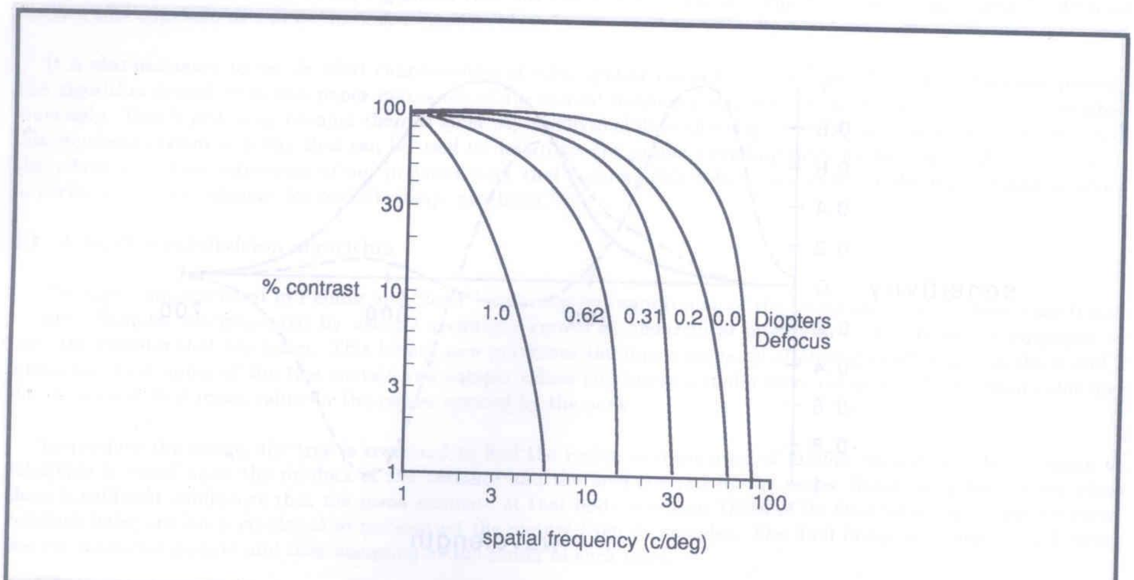

Figure 2: Spatial frequency filtering caused by different amounts of defocussing (from DeValois and DeValois[5]).

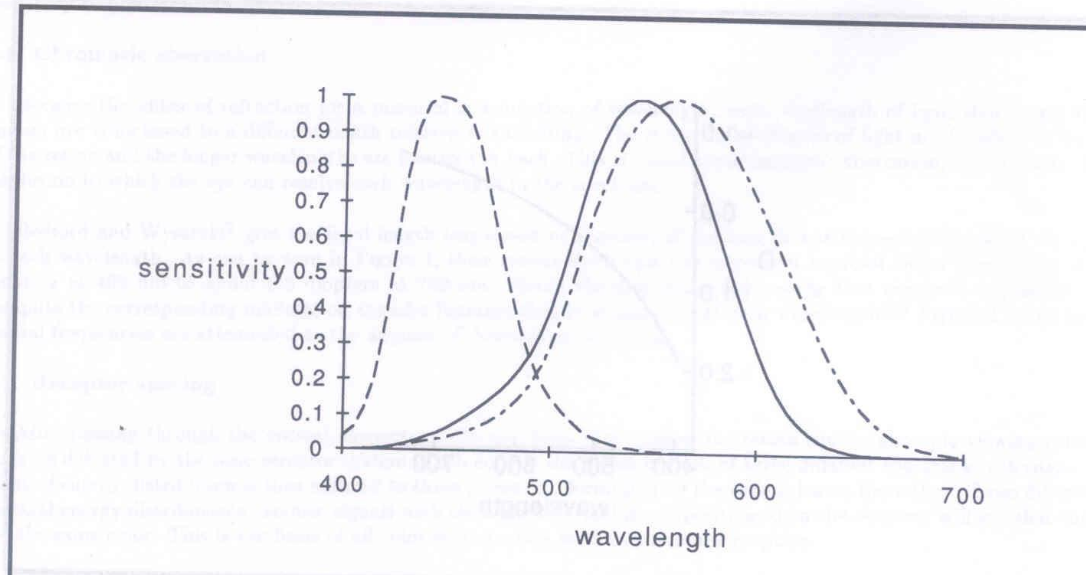

Figure 3: Spectral sensitivities for S (dashed), M (solid), and L (dot-dashed) components of SML space.

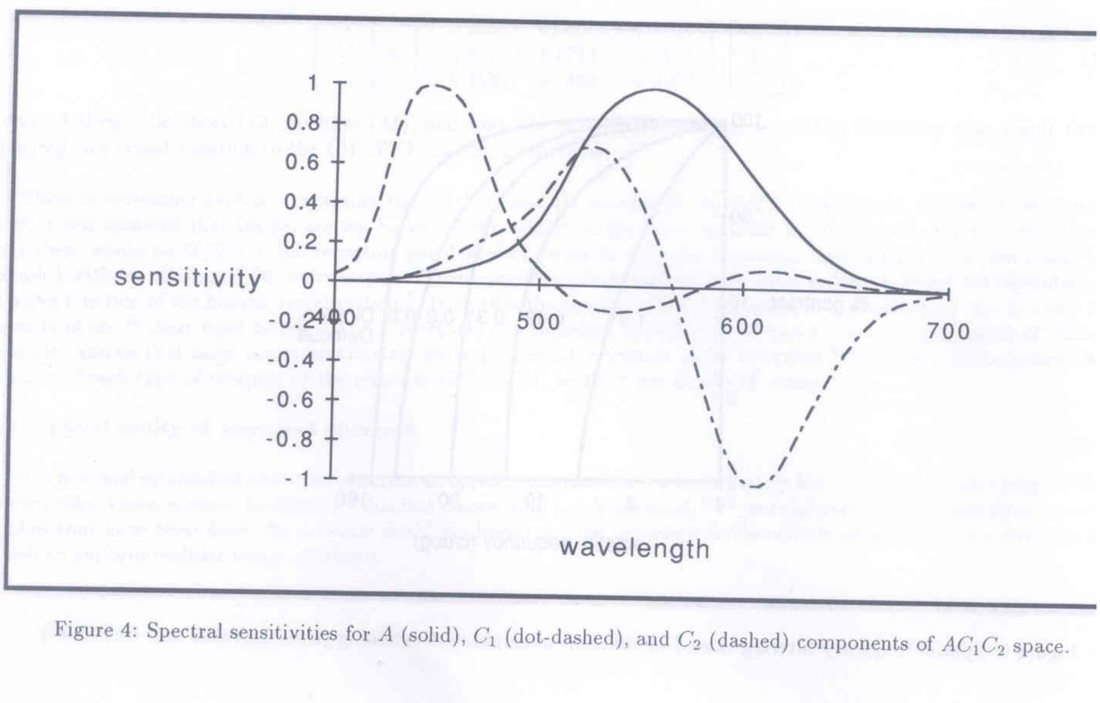

Figure 4: Spectral sensitivities for A (solid), C_1 (dot-dashed), and C_2 (dashed) components of AC_1C_2 space.

Most opponent representations for color are defined as a linear transformation of the SML receptor stage.[15,16] Given the above definition of SML space, this means that the opponent form can also be expressed as a linear transformation of the CIE XYZ color notation system. Buchsbaum and Gottschalk[14] have shown that when the discrete Karhunen-Loeve expansion is applied to the SML fundamentals, an opponent representation is achieved that is optimal from a statistical communication theory point of view and that is consistent with psychophysically derived opponent spaces. When this transformation is applied to the SML fundamentals defined in Equation 1 the following opponent color space is defined

$$\begin{bmatrix} A \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} -0.0177 & 1.0090 & 0.0073 \\ -1.5370 & 1.0821 & 0.3209 \\ 0.1946 & -0.2045 & 0.5264 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2}$$

Figure 4 shows the achromatic (A) and chromatic (C_1 and C_2) color channels that result.

Evidence has existed for some time that color spatial acuity is less than monochrome spatial acuity, and new evidence has recently strengthened this finding. Middleton[17] and Schade[18] did some of the earliest investigation of this subject and their work was employed to justify the use of a lower bandwidth for the the I and Q channels of the NTSC YIQ color transmission system. The YIQ system has since been shown to have characteristics similar to the opponents color processing stage of the human visual system.[19] Recently the work of Mullen[20] has provided specific figures regarding the spatial frequency bandpass of the achromatic and chromatic channels of the visual system. It was shown that cutoff for the luminance channel is 34 cycles/degree and for the red/green and yellow/blue channels it is 11 cycles/degree (Figure 5). The dimensions of the AC_1C_2 system are in good correspondence to the color space directions employed in Mullen.[20]

3. EXPLOITING HUMAN COLOR SPATIAL VISION

To take advantage of the limited color spatial acuity of the visual system it is necessary to employ an image synthesis algorithm that generates a picture by rendering low spatial frequency information first and high spatial frequency information last. In this way it is possible to control the amount of refinement that is used to produce the color spatial detail. Recently a screen subdivision algorithm that has this property was developed by Painter and Sloan.[21] We have adapted this approach as will be discussed below.

It is also necessary to decide what characteristic of color spatial vision to use in controlling the refinement process. The algorithm described in this paper makes use of the spatial frequency response of the AC_1C_2 color space described previously. This is primarily because there is good psychophysical data that expresses the spatial filtering properties of this opponent system in a way that can be used to determine the proper sampling rates for the three color channels. It also allows us to take advantage of our previous work that employs this color space to select the wavelengths at which to perform color calculations for realistic image synthesis.[1]

3.1. Adaptive subdivision algorithm

The algorithm described in Painter and Sloan[21] uses adaptive subdivision of the image plane to produce a ray traced picture. Samples are generated by using a technique known as hierarchical integration. A k-D tree is employed to store the samples that are taken. This binary tree partitions the image plane by alternately splitting it in the x and y directions. Leaf nodes of the tree contain raw sample values and internal nodes store estimates of the mean value and the variance of that mean value for the region covered by the node.

To produce the image, the tree is traversed to find the region in most need of further refinement. The decision to subdivide is based upon the product of the variance and the area covered by the node. Subdivision terminates when there is sufficient confidence that the mean estimate at that node is within 1/255 of its final value. A simple piecewise constant interpolation is employed to reconstruct the picture from its samples. The final image is produced by filtering the reconstructed picture and then sampling at the center of each pixel.

3.2. Modifications to the adaptive subdivision algorithm

The Painter and Sloan algorithm was modified to take advantage of the limited spatial acuity of the visual system's chromatic channels. The major change was to represent color in AC_1C_2 space instead of RGB space. Color was initially specified in terms of spectral reflectances and emittances, but the shading algorithm computed AC_1C_2 tristimulus values from the spectral energy distributions that resulted from tracing each ray. Images computed by the algorithm were stored as AC_1C_2 tristimulus values. This format is consistent with certain standards proposed for the storage and transmission of color images[19,22] but does require that the pictures be matrixed to RGB space before they can be displayed on a color television monitor.

Due to the restricted color spatial acuity of the visual system, the amount of image subdivision necessary to determine the C_1 and C_2 channels should be less than that required for the A channel. This means that while it is essential to traverse the entire depth of the k-D tree to determine the A component of the tristimulus values, the C_1 and C_2 components can be found by descending to a lesser depth of the tree. Leaves below a certain depth need only store the A component. This has the additional advantage of basing the test for termination on differences in luminance as recorded in the A component. This makes it easy to use brightness increments as the test for termination as suggested in Painter and Sloan although we did not do so in our implementation.

The result of using the algorithm to create a picture is illustrated in Figures 6 to 9. Figure 6 shows the three separate channels of the image, each computed to a high level of subdivision. (At the image resolution of 320 by 320 pixels shown, a 100K picture has been sampled at an average density of one ray per pixel.) The separate C_1 and C_2 channels in Figure 6 look somewhat like photographic negatives because the opponent functions for these channels produce negative tristimulus values (Figure 4) and the value stored for each pixel has been adjusted to accommodate this fact. What happens as the number of rays used to produce each channel of the image is varied is illustrated in Figure 7. Here the number of rays changes from 10K in the left picture to 50K in the right picture.

The effect of varying the resolution used to produce either the achromatic or chromatic channels is shown in Figures 8 and 9. In Figure 8 the number of rays used to produce the A channel has been varied from 10K in the left picture to 50K in the right picture while the number of rays used for the C_1 and C_2 channels has been held constant at 140K. In Figure 9 the situation has been reversed, with the number of rays used for the C_1 and C_2 channels changing from 10K to 50K while the number of rays for the A channel remained constant at 140K. Comparison of Figures 8 and 9 shows that reducing the number of rays used to create the C_1 and C_2 channels does not have as severe an effect on image quality as does a similar reduction of the rays used to create the A channel. A psychophysical study was performed in order to obtain a more quantitative statement regarding this effect. The results are discussed in Section 3.4.

3.3. Computational savings

The time complexity of a ray tracing algorithm is mainly composed of two parts: 1) the intersection calculations that are part of creating the ray tracing tree, and 2) the color shading calculations during traversal of the ray tracing tree to determine the intensity at each ray-surface intersection. For a given ray tracing tree, the time required for color calculations is approximately constant, but the time necessary for intersection calculations varies depending on the encountered object. There is little doubt that the time complexity of calculating intersections is dominant in ray tracing and several methods have been developed to minimize intersection calculations. However, the expense of computing color intensities cannot be neglected. In particular, when the objects in a ray traced scene are relatively simple the number of color calculations becomes increasingly significant. This paper is a first effort at minimizing these color calculations.

In the Painter and Sloan algorithm, determining the color produced by a given ray tracing tree requires that the shading calculations be performed for each of the tristimulus components. In our modified version, however, this work is only done for a tristimulus component when needed. As a way to save time the tristimulus components which are beyond the resolving power of the human eye at a certain stage of image subdivision are not calculated. If the constructed k-D tree is not well balanced, we will have a large time saving for color calculations. In this case, the samples will be taken at the areas where further refinement is most needed. If the stopping conditions for A, C_1, and C_2 are set at levels p, q, and r respectively, and the leaf nodes are distributed at all possible levels, then the total time saving will be at least $(2p - q - r)$ in contrast to the original cost $3p$. On the other hand, if the constructed k-D tree is well balanced because the samples taken are evenly distributed at the image plane, then screen subdivision will not be used efficiently nor will there be any significant time saving for color calculations.

There are also time savings in the reconstruction process. Before discussing this let us examine the reconstruct process described by Painter and Sloan. In this algorithm, the k-D tree generated in the refinement process partiti the image plane into different sized rectangular cells aligned with the x and y axes. Each rectangular cell contains sample point. In the reconstruction process, a color interpolation scheme is used to acquire a continuous image funct based on the k-D tree, then a low pass filter is used to resample the image function at the center of each pixel of screen. Because of our new approach to the early refinement process, the overall rectangular size varies from channel channel. As a result, there is a moderate efficiency improvement in the reconstruction process. This can be exemplif in the case of a piecewise constant interpolation. The C_1 and C_2 channels have overall bigger size rectangular cells comparison to the A channel. Thus it takes less time to interpolate the color values between samples in addition requiring a shallower traversal of the k-D tree.

3.4. Psychophysical testing

A psychophysical test was conducted to better evaluate the new screen subdivision ray tracer. The hypothesis to tested was that reducing the number of rays used to synthesize the chromatic channels of a picture has less of an eff on image quality than a similar reduction in the number of rays used to produce the achromatic channel. This is t result that appears to be illustrated by the sequence of images in Figures 8 and 9.

To test this hypothesis, two sequences of the same image were synthesized. In one sequence the number of rays us to produce the achromatic channel (A) was varied from 10K to 140K in steps of 10K while the number of rays us to produce the chromatic channels (C_1 and C_2) was held constant at 140K. In the second sequence the number of ra for the chromatic channels changed from 10K to 140K while the number of rays used for achromatic channel stayed 140K. All images were created at a resolution of 320 by 320 pixels. This resolution creates an average sampling dens of greater than one ray per pixel whenever the total number of rays exceeds approximately 100K, and it insures that t sampling rate will be sufficient for the highest detectable spatial frequencies when the image is viewed from the distar used in the experiment (see below). Figures 8 and 9 show some of the images in the sequences that were used. Re green, yellow, blue, white and black (in the shadows) were employed in the image to insure that edges with colors fro the poles of the opponent space were present.

Subjects were asked to evaluate the quality of each of the images in these two sequences. A method similar to th employed in Atherton and Caporael,[23] Booth, et. al.,[24] and Sandford, et. al.[25] was used. Each image was present randomly to the subject and they were asked to assign it a numerical rating between zero and ten. To provide a ba for this judgement, an image corresponding to a rating of zero was presented to the left of the test image and an ima corresponding to a rating of ten was presented to the right of the test image. The zero quality image was created using 10K rays for both the achromatic and chromatic channels, and the ten quality image was produced by using 14C rays for both types of channels. These anchor images remained constant throughout the test.

Subjects that participated in the test were primarily graduate students in the Department of Computer and Info mation Science at the University of Oregon and almost none of them had a background in computer graphics. The te stimuli were presented on a Barco Calibrator monitor that was balanced to produce D6500 white whenever $R = G =$ Subjects were seated 40 inches from the monitor and were instructed not to lean forward to get closer to the display. (this distance each pixel subtended approximately 0.019 degrees of vision.) Each test image was presented for ten secon after which the subject used a mouse to select a rating. It took each subject approximately 20 minutes to complete tl experiment.

The mean results for 12 subjects are plotted in Figure 10. It is clear that image quality decreases with the number rays used to create the achromatic and chromatic channels, and an analysis of variance shows that this effect is significa ($F(12,120)=51.44$, $p < 0.001$). Adjusting the achromatic or chromatic channels also had a significant effect on the resul ($F(1,10)=24.03$, $p < 0.001$) with changes to the chromatic channels having less of an impact on image quality at a choices for number of rays. In addition, there was an interaction between the two variables ($F(12,120)=11.62$, $p < 0.001$

As the number of rays increases beyond a certain point, the ratings for the both the achromatic and chromatic case approach an asymptote. Beyond this point adding more rays to the image has little or no effect on the quality of tl

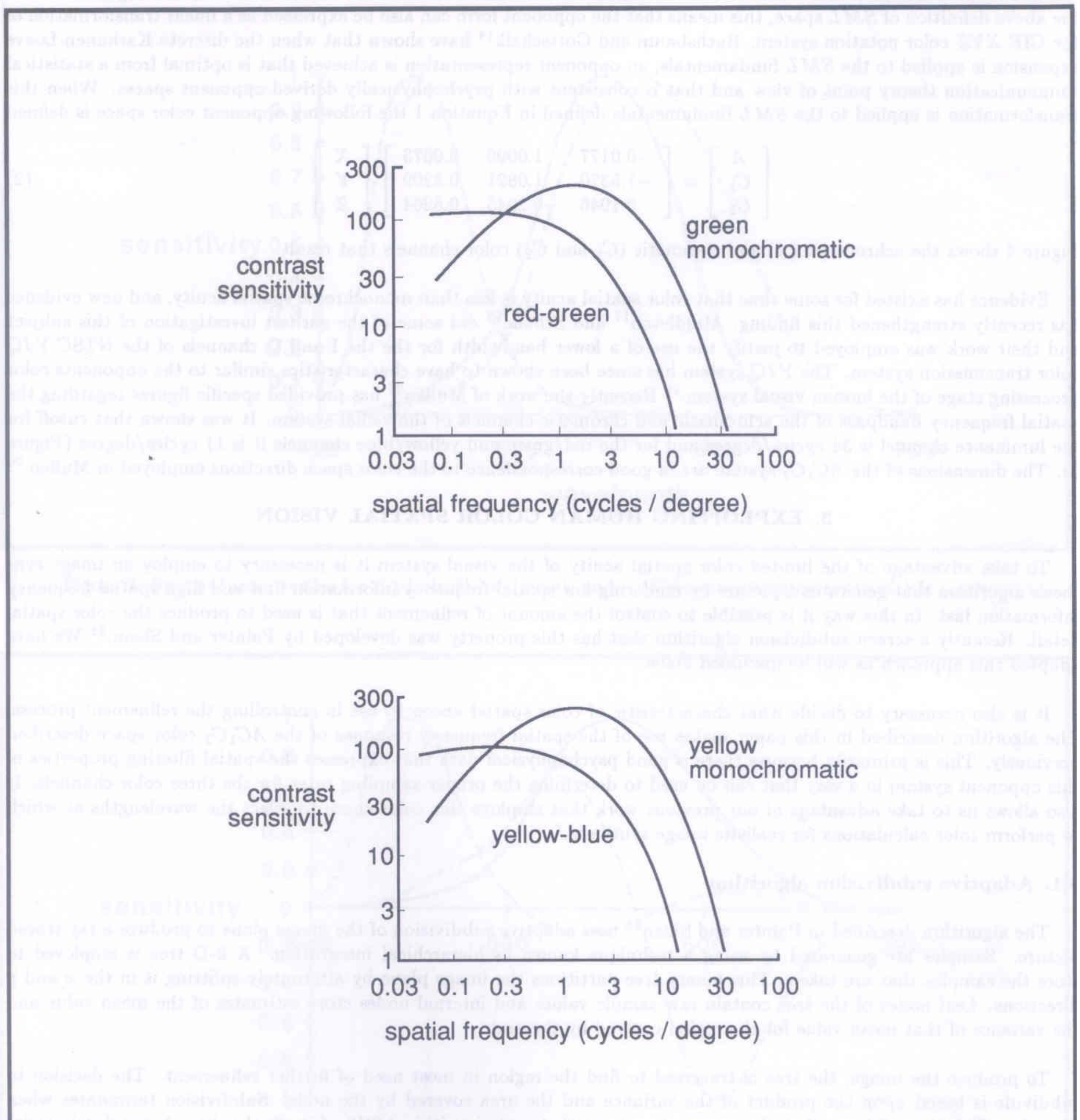

Figure 5: Spatial frequency response of the opponent color channels. Red-green versus monochromatic green in the top diagram and blue-yellow versus monochromatic yellow in the bottom diagram (from Mullen[20]).

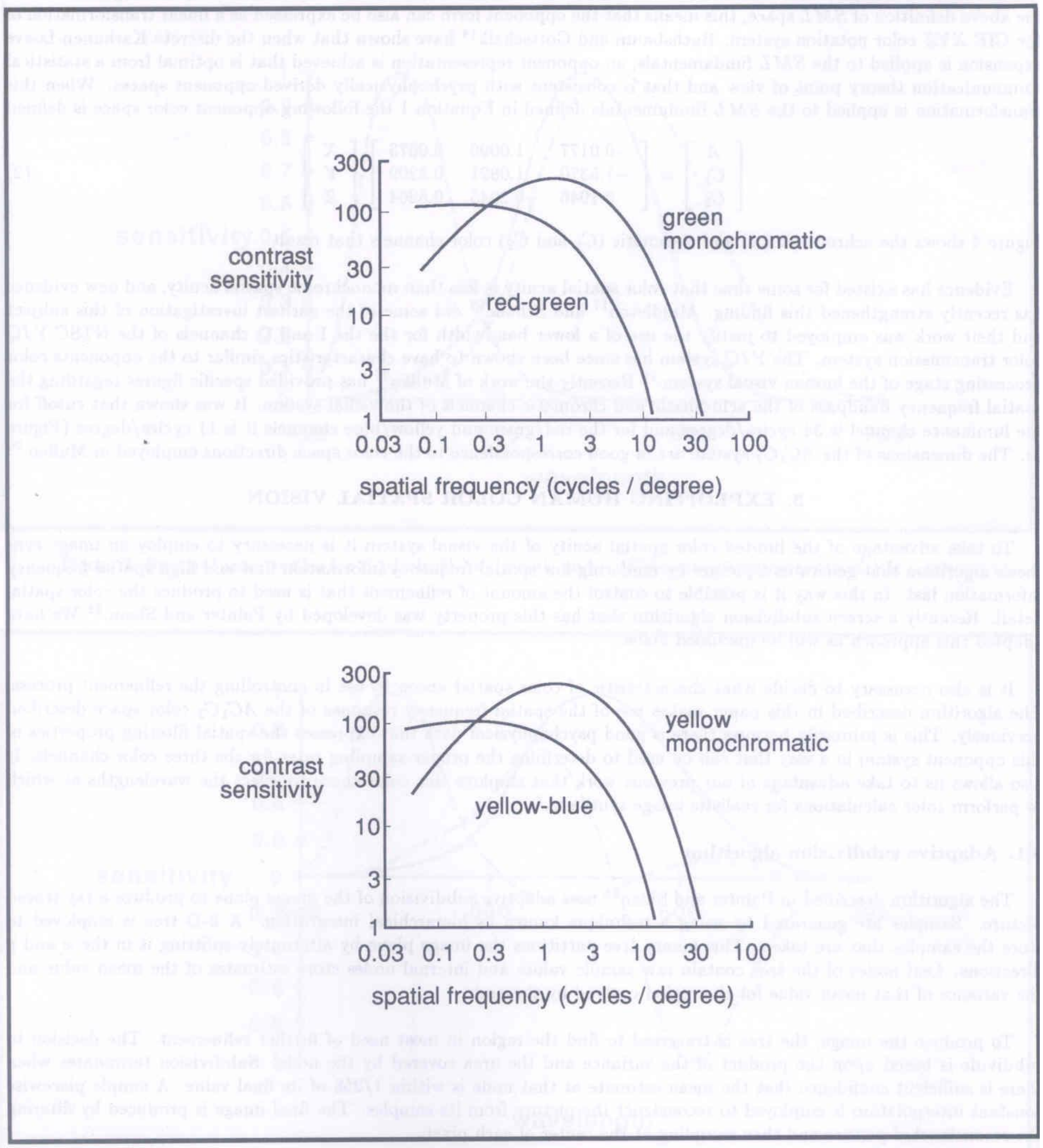

Figure 5: Spatial frequency response of the opponent color channels. Red-green versus monochromatic green in the top diagram and blue-yellow versus monochromatic yellow in the bottom diagram (from Mullen[20]).

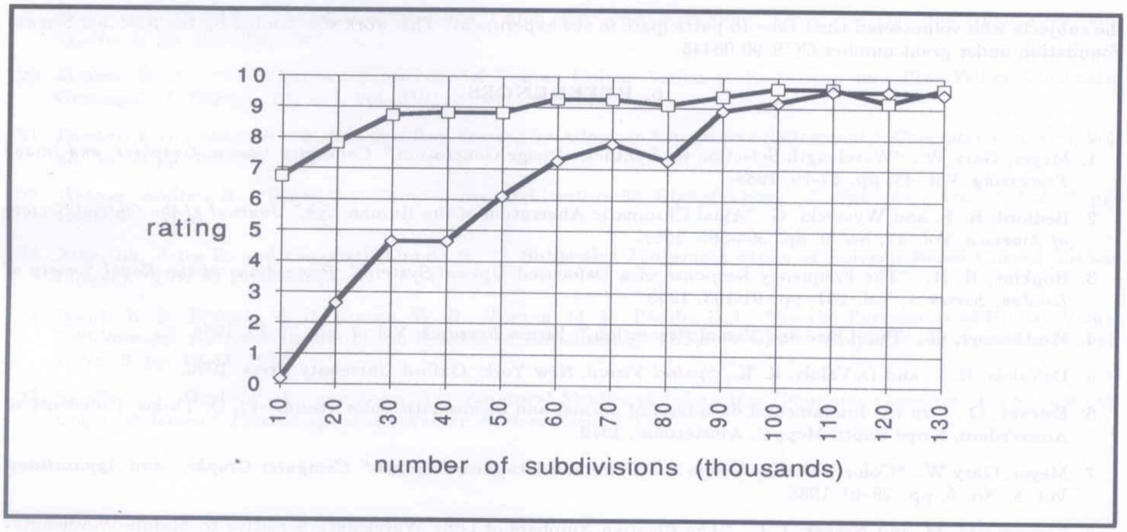

Figure 10: Image quality rating versus number of spatial samples used to produce chrominance information (squares) and luminance information (diamonds).

image. While it is difficult to say with certainty where each of these two curves flattens out, it does appear that major changes occur in the ratings of the luminance adjusted images at about 90K and in the ratings of the chrominance adjusted images at about 30K.

4. SUMMARY AND CONCLUSIONS

A realistic image synthesis algorithm was developed that takes advantage of the limited color spatial acuity of the human visual system. A screen subdivision ray tracer was written that limits the depth to which a k-D tree must be descended in order to compute the C_1 and C_2 components of AC_1C_2 tristimulus values. Psychophysical results regarding the spatial frequency response of the C_1 and C_2 channels were used to direct the depth to which the tree was traversed. A modest savings in computational effort was realized and perceptual tests showed that decreasing the number of rays used to produce the chromatic channels had less of an effect on image quality than reducing the number of rays used to create the achromatic channels.

Perhaps the most important contribution of this work is that it shows how the characteristics of the visual perceptual system can be used to facilitate the production of realistic images. It also represents a first attempt to minimize the cost of performing the color calculations in a ray tracer just as others have worked at decreasing the costly object intersection computations. The use of an opponents color space dovetails nicely with previous work that we have done regarding wavelength selection and it is consistent with proposals that have been made for using a perceptual components approach for storing and transmitting images. It also provides a platform, solidly grounded in what is known about color psychophysics, from which other perceptually based improvements to realistic image synthesis can be explored.

5. ACKNOWLEDGEMENTS

Mark VandeWettering did the initial implementation of the algorithm described in Painter and Sloan.[21] Michael Kelly wrote the software that was used to perform the perceptual comparison tests. The authors would like to thank the subjects who volunteered their time to participate in the experiment. This work was funded by the National Science Foundation under grant number CCR 90-08445.

6. REFERENCES

1. Meyer, Gary W., "Wavelength Selection for Synthetic Image Generation," *Computer Vision, Graphics, and Image Processing*, Vol. 41, pp. 57-79, 1988.

2. Bedford, R. E. and Wyszecki, G. "Axial Chromatic Aberration of the Human Eye," *Journal of the Optical Society of America*, Vol. 47, No. 6, pp. 564-565, 1957.

3. Hopkins, H. H., "The Frequency Response of a Defocused Optical System," *Proceedings of the Royal Society of London, Series A*, Vol. 231, pp. 91-103, 1955.

4. Westheimer, G., "Pupil Size and Visual Resolution," *Vision Research*, Vol. 4, pp. 39-45, 1964.

5. DeValois, R. L. and DeValois, K. K., *Spatial Vision*, New York: Oxford University Press, 1990.

6. Estevez, O., "On the fundamental data-base of normal and dichromatic color vision," Ph.D. Thesis, University of Amsterdam, Krips Repro Meppel, Amsterdam, 1979.

7. Meyer, Gary W., "Color-Defective Vision and Computer Graphics Displays," *Computer Graphics and Applications*, Vol. 8, No. 5, pp. 28-40, 1988.

8. Cicerone, C. M. and Nerger, J. L., "The Relative Numbers of Long-Wavelength-Sensitive to Middle-Wavelength-Sensitive Cones in the Human Fovea Centralis," *Vision Research*, Vol. 29, No. 1, pp. 115-128, 1989.

9. Campbell, F. W. and Robson, J. G., "Application of Fourier Analysis to the Visibility of Gratings," *Journal of Physiology*, Vol. 197, pp. 551-566, 1968.

10. Williams, D. R., MacLeod, D. I. A., and Hayhoe, M. M., "Punctuate Sensitivity of the Blue-Sensitive Mechanism," *Vision Research*, Vol. 21, pp. 1357-1375, 1981.

11. Jameson, D. and Hurvich, L. M., "Some Quantitative Aspects of an Opponent-Colors Theory, I. Chromatic Responses and Spectral Saturation," *Journal of the Optical Society of America*, Vol. 45, pp. 546-552, 1955.

12. Hurvich, L. M. and Jameson, D., "Some Quantitative Aspects of an Opponent-Colors Theory, II. Brightness, Saturation, and Hue in Normal and Dichromatic Vision," *Journal of the Optical Society of America*, Vol. 45, pp. 602-616, 1955.

13. DeValois, R. L., Smith, C. J., Karoly, A. J., and Kitai, S. T., "Electrical Responses of Primate Visual System, I. Different Layers of Macaque Lateral Geniculate Nucleus," *Journal of Comparative and Physiological Psychology*, Vol. 51, pp. 662-668, 1958.

14. Buchsbaum, G. and Gottschalk, A. "Trichromacy, Opponent Colours Coding and Optimum Colour Information Transmission in the Retina," *Proceedings of the Royal Society of London, Series B*, Vol. 220, pp. 89-113, 1983.

15. Ingling, C. R., "The Spectral Sensitivity of the Opponent-Colors Channels," *Vision Research*, Vol. 17 pp. 1083-1090, 1977.

16. Guth, S. L., Massof, R. W., and Benzschawel, T., "Vector Model for Normal and Dichromatic Color Vision," *J. Opt. Soc. Am.*, Vol. 70, pp. 197-211, 1980.

17. Middleton, W. E. and Holmes, M. C., "The Apparent Colors of Surfaces of Small Subtense - A Preliminary Report," *Journal of the Optical Society of America*, Vol. 39, pp. 582-592, 1949.

18. Schade, Otto H., "On the Quality of Color-Television Images and the Perception of Color Detail," *Journal of the Society of Motion Picture and Television Engineers*, Vol. 67, No. 12, pp. 801-819, 1958.

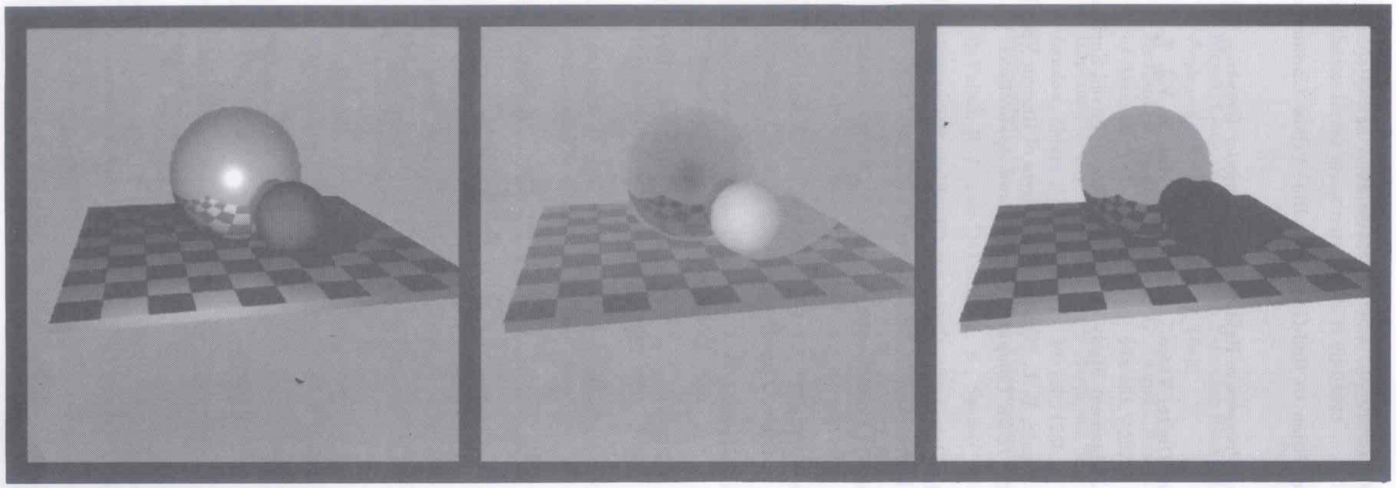

Figure 6: The A, C_1, and C_2 channels (left to right) of the test image. 140K spatial samples were used to produce all three component images. (Inverted grayscale in C_1 image caused by negative values in C_1 channel.)

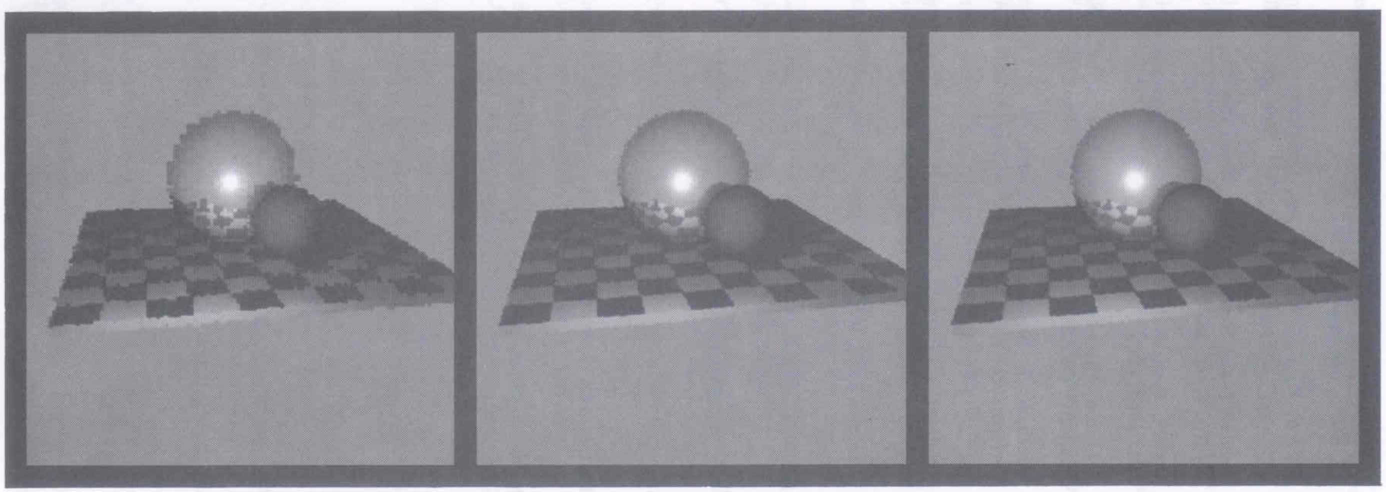

Figure 7: Result of using 10K, 30K, and 50K spatial samples (left to right) to produce the A channel of the test image.

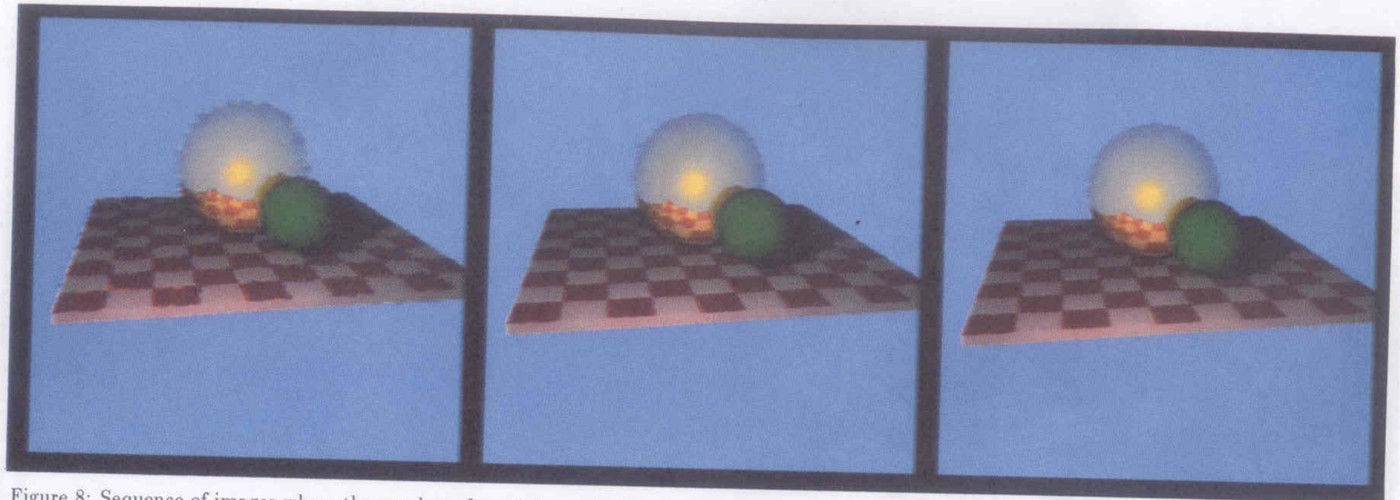

Figure 8: Sequence of images where the number of spatial samples for the A channel varies as shown in Figure 7 while the number of samples for the C_1 and C_2 channels remains constant at 140K.

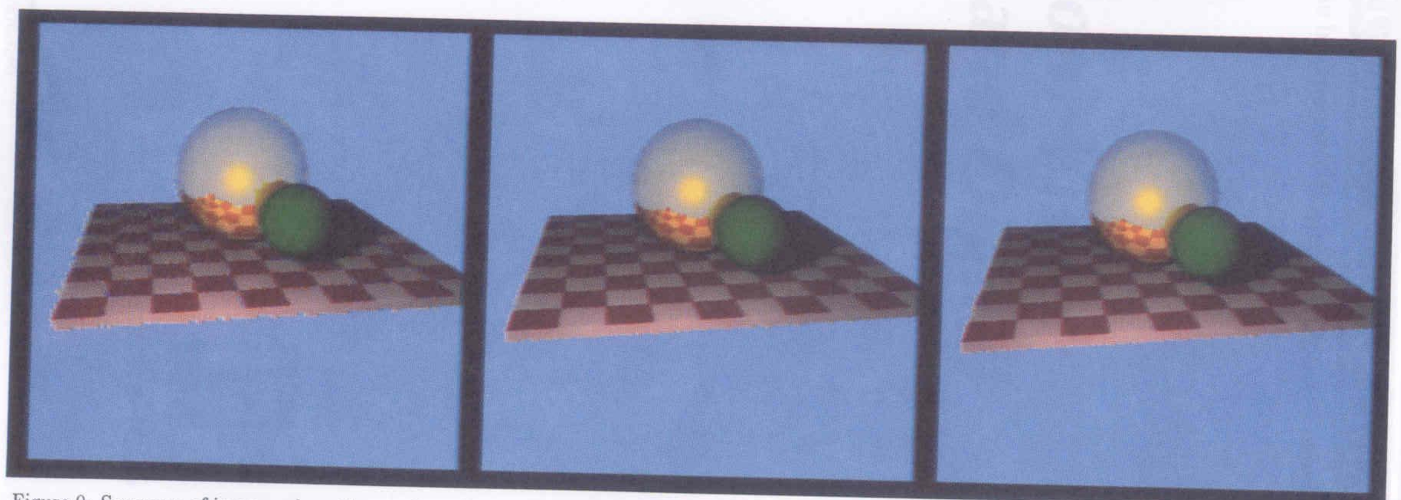

Figure 9: Sequence of images where the number of spatial samples for the C_1 and C_2 channels varies as shown in Figure 7 while the number of samples for the A channel remains constant at 140K.