Annals of the Royal College of Surgeons of England (1975) *vol* 57

Colour television, an imitation of the human visual system

R W G Hunt DSC

Research Division, Kodak Limited, Harrow, Middlesex

Summary

Colour television is examined as an attempt to imitate the human visual system in image formation, spectral sensitivities, adaptation, contrast effects, signal processing, signal modulation, signal transmission, and interpretation. Many similarities are found, but where differences occur they are discussed in relation to the differences between visual and television technologies.

Introduction

At first sight colour television and human colour vision might seem to have little in common, the former involving rather massive equipment containing copper wires, silicon transistors, and evacuated glass tubes while the latter depends entirely on biological material. But although the materials used are so different in the two cases, colour television succeeds only because it imitates human vision. It is convenient to examine this imitation by dividing the visual system into eight successive steps: (1) image formation; (2) spectral sensitivities; (3) adaptation; (4) contrast effects; (5) signal processing; (6) signal modulation; (7) signal transmission; and (8) interpretation.

Image formation

The eye, with its extremely wide field of view, selects those parts of a scene having special interest and examines their details one at a time; this is done by movements of the eyes to bring each detail on to the fovea, where visual acuity is greatest. But in television acuity is the same all over the picture and is limited by the number of lines used in the system. The 'select and examine in detail' stratagem of the eye has therefore to be imitated in television by the use of several cameras (usually three) with lenses of different focal lengths to provide, successively, general views, close-up views, and extreme close-ups; the use of zoom lenses on each camera further aids in the 'select and examine' procedure.

It has been shown that when light falls on the retina of the eye it is absorbed by a mosaic of light-sensitive cells of four different types: rods, 'red-absorbing' cones, 'green-absorbing' cones, and 'blue-absorbing' cones[1]. The rods provide a highly sensitive monochrome facility on which vision depends at low light levels and the cones give us colour vision at normal and high light levels. This highly convenient monochrome–colour arrangement is imitated in television only by the use of separate cameras. The colour part of the retina is imitated by splitting the light from the camera lens into reddish, greenish, and bluish components by means of dichroic interference semireflecting mirrors.

Edridge-Green Lecture delivered on 13 June 1974, accompanied by many visual illustrations.

Although this avoids wasting light, it means that colour cameras have to be very precisely constructed to ensure that the three images are scanned in exact registration, a feature which, with others, results in colour cameras costing around 10 times as much as monochrome cameras. A mosaic type of camera, imitating the visual arrangement of red-, green-, and blue-sensitive cells situated side by side, although optically less efficient, would have many advantages if it could be manufactured; the charge-coupled-device arrays now being produced may eventually make this possible. Meanwhile much ingenuity has been exercised to make the beam-splitting cameras self-correcting for registration[2], and this has much improved their reliability, although at increased cost.

Spectral sensitivities

The spectral sensitivities of the three different types of retinal cone are approximately as shown in Fig. 1 by the curves marked ρ, γ, and β (for the 'red-absorbing', 'green-absorbing' and 'blue-absorbing' cones respectively). Also shown in this figure are the spectral emission curves of red, green, and blue phosphors typical of those of modern colour television display tubes. Ideally, each phosphor should emit light that stimulates only one of the three types of cone: red light for the ρ-cones, green light for the γ-cones, and blue light for the β-cones. Although this is approximately the case for the red and blue phosphors, the green phosphor results in substantial stimulation of the ρ-cones. (This could be avoided only by using a green phosphor emitting at shorter wavelengths, but then substantial stimulation of the β-cones would occur; the difficulty is a fundamental one, caused by the overlapping of the ρ and β curves leaving no wavelengths at which the γ-cones alone have sensitivity.)

Hence if a television camera imitated the ρ, γ, and β sensitivity curves of the eye, although the signals in the red, green, and blue channels would be correct, the light emitted from the green phosphor would produce an unwanted response in the ρ-cones as well as the wanted response in the γ-cones.

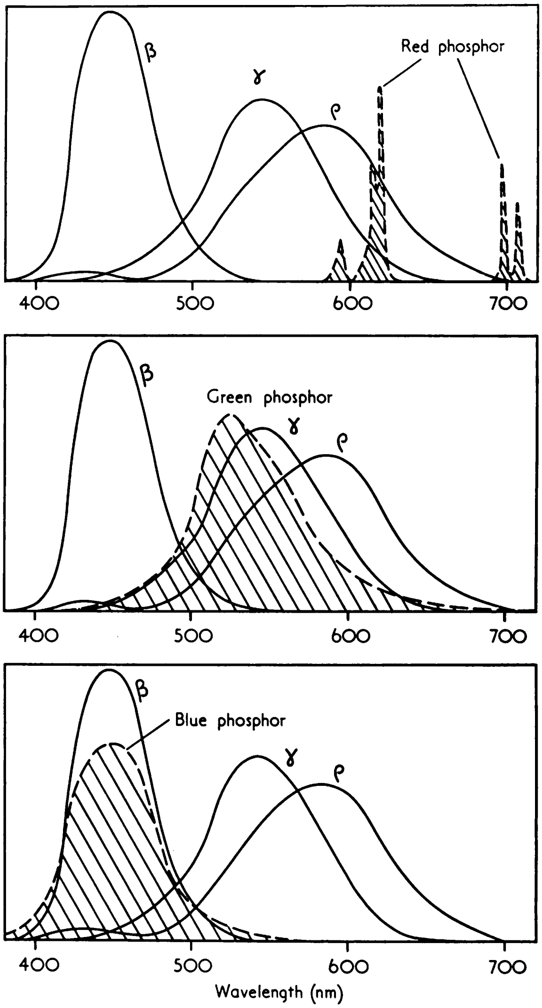

FIG. 1 *Probable spectral sensitivities, ρ, γ, and β, of the three different types of retinal cone together with spectral power distributions typical of the red-, green-, and blue-emitting phosphors used in colour television receivers.*

It is, however, possible to allow for this by arranging for the red television signal to have subtracted from it a correction signal equal to a suitable fraction of the green signal. If the correction signal is greater than the red signal the resultant signal becomes negative, and since negative light cannot be produced, errors result; but for most colours the correction signal is less than the red signal. This procedure is known as 'matrixing' and is usually elaborated to correct for the unwanted stimulations produced by all three phosphors. In this case it is no longer necessary to have the ρ, γ, β sensitivity curves; any three linear combinations of the three curves can be used instead, provided that the matrixing is suitably adjusted. The camera sensitivity curves used in practice are usually chosen to overlap at their 50% levels so as to avoid wasting light in the beam-splitter; and the display phosphors need to produce as little unwanted cone stimulation as possible so as to minimize the number of colours for which one of the signals becomes negative (because such colours cannot be reproduced by the system).

Adaptation

The ability of the eye to adjust its sensitivity over wide ranges of illumination levels is well known, but it can also adjust the relative sensitivities of its three types of retinal cone so as to compensate very largely for changes in the colour of various illuminants. Television cameras therefore have to have facilities for adjusting not only the overall illumination level (such as variable lens apertures) but also for illuminant colour, and this is usually done by using separate gain controls on the three channels to achieve equality of the three signals for a reference white or grey in the scene.

As the level of illumination of a scene is raised, the apparent vividness of the colours increases. This means that the higher the luminous intensity of a television display, the more vivid will be its colours. However, phosphors of high luminous efficiency tend to be those emitting light mainly near the centre of the visible spectrum, and these tend to produce large unwanted cone stimulations. A compromise is therefore necessary in choosing phosphors, the aim being to produce a picture of reasonably high luminance with as few colours as possible being unreproducible on account of one of the signals being negative.

Contrast effects

Television pictures are usually viewed with a surround of appreciably lower luminance than that of the picture. It has been found that this dim surround causes the apparent contrast of the picture to be lower than if it is viewed with a surround of luminance similar to that of the average for the picture. It is therefore necessary for the contrast actually displayed to be higher than that of the original scene. The magnitude of the effect is illustrated in Fig. 2, in which reproduced (optical) density is plotted against log scene luminance (relative to a suitable reference white). The slope of the line relating these two quantities, the gamma, is unity only if the picture has a surround of luminance similar to that of the average for the picture ('average surround'); for dim surrounds the gamma required is about 1.25 and for dark surrounds (such as occur when, for instance, a film is projected in a darkened auditorium) the gamma required is about 1.5[3]. (The presence of flare in reproduction systems and certain other considerations result in the gamma required in colour reproduction systems in practice being different at different density levels[4].)

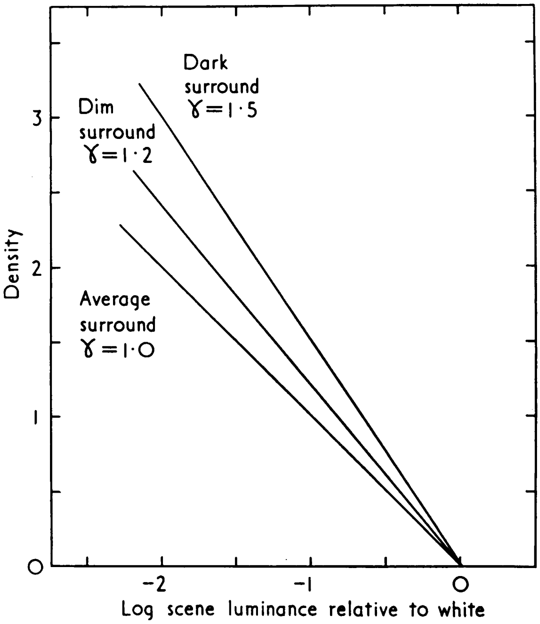

FIG. 2 *Density required in reproduction for dark surround, dim surround, and 'average' surround (luminance of surround equal to that of the average of the picture) conditions plotted against log scene luminance relative to white.*

Signal processing

It has long been suggested that the signals sent from the retina to the brain are not functions of the absorptions in the three types of retinal cone indicating 'redness', 'greenness', and 'blueness' but functions of combinations of these absorptions indicating lightness or darkness, redness or greenness, and blueness or yellowness. Physiological studies in the lateral geniculate body of the macaque monkey have lent strong weight to this view being correct[5].

When colour television signals are transmitted this visual coding of the signals is closely imitated. Thus instead of red, green, and blue signals a luminance signal, a red-minus-luminance colour-difference signal, and a blue-minus-luminance colour-difference signal are transmitted. It is found that by doing this the total amount of information required to be transmitted is much reduced because the two colour-difference signals need contain only one-sixteenth of the amount of information per unit area as the luminance signal. This follows from the fact that more cones are required to signal correctly a change in the visual colour-difference signals than a change in the visual lightness–darkness signal[6].

Signal modulation

Although a few workers[7] have reported colour-dependent time distributions of nerve impulses, it is generally held that visual signals are modulated on a pulse-frequency basis—that is, the strength of the signal depends on the rate of nerve impulses per unit time. In colour television amplitude modulation of sinusoidal signals is most widely used, but frequency modulation of sinusoidal signals is also used to some extent (for instance, in the SECAM system for the colour-difference signals, and generally for videotape recording). Digital signals, with pulse-code modulation, are used in converting television signals from one standard to another (for example, 625-line pictures to 525-line pictures). So far television has not made use of the pulse-frequency modulating method used in vision, but frequency modulation and pulse-code modulation may perhaps be regarded as variations of the visual method.

Signal transmission

The use of two eyes and the pooling of information from both at the optic chiasma provide the basis for stereoscopic vision. So far no attempt has been made to provide stereoscopic colour television and it seems unlikely that this will be done. Even in photography, where the difficulties are not so great, little use has been made of stereoscopic pictures. The reason is probably that stereoscopic vision is of practical importance only for objects within a distance of about a metre, and most pictures are mainly concerned with objects at greater distances. Moreover, the solidity of objects and their relative positions in space can be very effectively conveyed by good lighting and careful choice of camera angles.

Interpretation

A vital part of the visual system is the brain of the observer, and it is fortunate for colour television that the final processes occur in the brain of the viewer. The television receiver produces patches of coloured light and the viewer interprets these as objects with which he is familiar. However, when the original scene is outside the viewer's experience—as in some pictures from space, for instance—it may not then be possible to perceive the scene meaningfully. The brain thus plays an important part, and good television avoids visual ambiguities, an example of which is given in Fig. 3.

Conclusion

Colour television is a marvel of modern technology—high-quality colour pictures produced simultaneously in millions of homes, transcontinental reception by satellite, and even pictures from the moon. And yet, in a sense, when compared with the visual system colour television still appears perhaps a little crude. The beam-splitting colour camera, the large evacuated cathode-ray tube in the receiver, the laborious repetition of nearly identical information on successive television fields compare strangely with the all-solid-state colour mosaic of the retina and the exception-reporting select-and-examine stratagem of vision. But comparison of the two systems is surely instructive in our understanding of both, and had Edridge-Green been active today I am sure he would have been foremost amongst those involved in such a study.

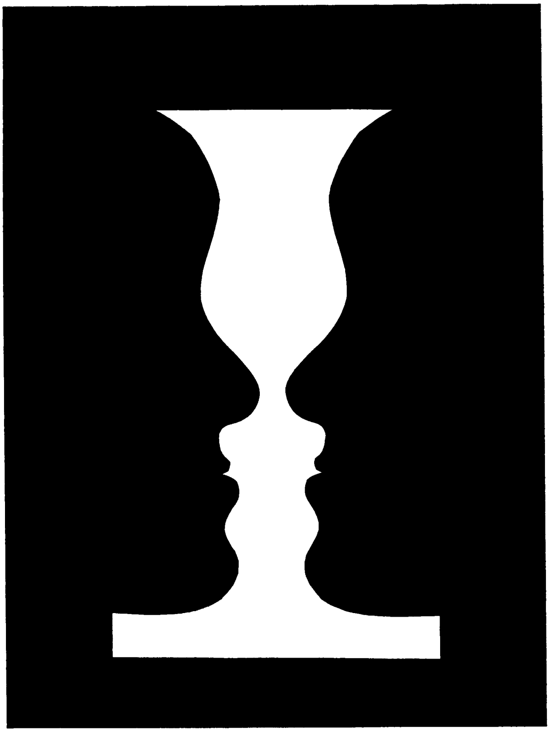

FIG. 3 *The interpretation of the visual pattern by the brain is an important part of the perceptual process. In this example the same visual image can result in the perception either of two faces looking at one another or of a white vase on a black background.*

References

1 Marks, W B, Dobelle, W H, and MacNichol, E F (1964) *Science,* 143, 1181.

2 Underhill, W T (1971) *Journal of the Royal Television Society,* 13, 191.

3 Bartleson, C J, and Breneman, E J (1967) *Photographic Science and Engineering,* 11, 254 and *Journal of the Optical Society of America,* 57, 953.

4 Hunt, R W G (1969) *British Kinematography, Sound and Television,* 51, 268.

5 DeValois, R L (1970) *Proceedings of the 1st AIC Congress,* Stockholm, p. 29.

6 Hunt, R W G (1967) *Journal of the Royal Television Society,* 11, 220.

7 Kozak, W M, and Reitboeck, H J (1974) *Vision Research,* 14, 405.