This paper appeared in the January, 1999 issue of the Proceedings of the IEEE

Introduction to "Color Television—Part I"

WILLIAM F. SCHREIBER, LIFE FELLOW, IEEE, AND ROBERT R. BUCKLEY

Invited Paper

The field-sequential color-TV system discussed in this Classic Paper was the first such system that actually worked well and produced good-quality pictures. It suffered from inadequate spatial and temporal resolution due to the need to send three pictures in the channel that was designed for one. Its use of the spinning color wheel was the subject of much derision, but the main problem with the system was inefficient use of bandwidth due to failure to take into account well-known principles of color vision. The system was adopted by the Federal Communications Commission (FCC) in 1950 as the U.S. standard, but the industry boycotted it and it was replaced in 1953 by the National Television Systems Committee (NTSC) system, still in use.

For a better understanding of the milieu in which this system was developed, we first present a summary of color reproduction science and systems as of 1942. The main contribution of Goldmark et al. was the design of a workable manufacturable system using existing mechanical and electronic components available at the time, which system approximated the performance known to be required for good color reproduction. We then analyze the paper itself and find that it was typical of papers of the time, which tended to go into great detail about construction of circuits but failed to analyze the central problem in sufficient depth.

Finally, we give a brief description of the progress of color TV since 1942 and try to draw some useful parallels between the color controversy of the time and the current arguments related to the transition to digital TV broadcasting.

***Keywords*—** *Color television, FCC standards decisions, field-sequential, television standards.*

I. PROLOGUE

Color television, today's most important source of news and entertainment, saturates society to an extent that makes it difficult to think of what life must have been like without it, or without television entirely. Yet it was not that long ago, within the memory of many still living, that black-and-white TV, standardized in 1941 and delayed only on account of World War II, had not yet come to the market. Color TV was assumed to be much further in the future when Goldmark and his colleagues at Columbia Broadcasting System (CBS) succeeded in demonstrating remarkably good color pictures using quite practical equipment which was readily manufacturable at that time. Moreover, hardly any adjustments were required at the receiver to maintain the picture quality. This is, therefore, a landmark paper that is definitely worth looking at again to review the early history of what has turned out to be a large and powerful industry, to note the differences between technology development then and now, and perhaps even to learn some lessons that can be applied to current technological controversies.

As it turned out, the field-sequential technique that was the basis of the CBS system was a dead end. Although a later version of the system was adopted by the Federal Communications System in 1950 as the U.S. color standard, work was already well underway by then on the so-called "simultaneous" system that did not use the rotating filter wheel that was such an object of derision. The CBS system was boycotted by virtually the entire industry, and in 1953 the Federal Communications Commission (FCC) reversed itself and adopted the National Television Systems Committee (NTSC) color system. Nevertheless, it is still worth looking at the genuine accomplishments of the field-sequential system and especially pondering the techno-political battle that developed between the two approaches. If that episode had been better understood by more of today's actors, the current controversy over the conversion to digital television transmission might have followed a very different course.

II. COLOR REPRODUCTION BEFORE GOLDMARK'S CLASSIC PAPER

The art of three-color reproduction can be traced to LeBlon in the early eighteenth century. After trying to make engraved color prints using the seven colors of Newton, LeBlon discovered that he could obtain a very large range of colors by mixing only three colors. A corresponding three-color theory of vision and its key corollary, showing that the appearance of almost any color can be produced by additive mixing of three primary colored lights, dates to Young in the first decade of the nineteenth century, although

---

Manuscript received September 4, 1998.

W. F. Schreiber is with Massachusetts Institute of Technology, Cambridge, MA 02139 USA

R. Buckley is with Architecture and Document Services, Xerox Corporation, Webster, NY 14580 USA.

Publisher Item Identifier S 0018-9219(99)00433-8.

Palmer had postulated three visual receptors for color 25 years before Young.[1]

In the 1850's and 1860's, Maxwell demonstrated additive color reproduction by successively taking black-and-white records of a color scene through red, green, and blue filters and then simultaneously projecting them in register through the same red, green, and blue filters to obtain the color rendition of the scene. Maxwell was also the first, followed by others (most notably Helmholtz), to calculate color mixture curves—in effect, the spectral taking sensitivities of the human visual system. In 1900, Ives described the relationship between the colored lights or primaries used to synthesize a color image and the filters used to analyze or make the corresponding color records of the original scene. Precise knowledge of the derivation of the color mixture curves, which is essential for accurate color rendition, had to wait for the publication of the first standard data, based on extensive psychophysical testing, by the International Commission on Illumination (CIE) in 1931. This put three-color reproduction on a sound theoretical and experimental foundation.

The era of numerical colorimetry (the science of color measurement) and accurate visual color matching based on the three-color theory thus began in 1931. In this system, now used universally, a color image is reproduced as three monochrome images, each controlling the light output of a different light source (display primary); the primaries are roughly red, green, and blue. The black-and-white images corresponding to the three primaries comprise a point-by-point weighted measure, called the tristimulus values, of the radiant power in a different spectral region (camera taking sensitivities).

The choice of display primaries dictates the three taking sensitivities for producing tristimulus values and obtaining accurate color rendition. Under these conditions, the reproduced color image visually matches the original scene, assuming proportionality of light intensities through the system. As Ives[2] had earlier observed, the choice of display primaries constrains the camera taking sensitivities. Hardy and Wurzburg supplied the mathematical formulation of this constraint in 1937, using the 1931 CIE color mixture curves.

In the 1931 CIE system of colorimetry, the nonnegative $\bar{x}$, $\bar{y}$, $\bar{z}$ color-mixture curves are taking sensitivities corresponding to imaginary display primaries that cannot be synthesized from any combination of positive amounts of real spectrum colors. Real, nonnegative display primaries require taking sensitivities with negative values at some wavelengths, and vice versa. The reason for this is that there are no primaries that both enclose and are enclosed by the horseshoe-shaped spectrum locus and purple line that contains all physical stimuli on the CIE chromaticity diagram. Since real display colors must have positive spectral power at all wavelengths, the corresponding camera taking sensitivities must have negative lobes, an example being given in Fig. 4 of the Goldmark paper.

Such taking sensitivities are physically unattainable, at least with ordinary color filters. All-positive sensitivities must be used, which causes erroneous color rendition. These errors can be reduced and even eliminated by processing electronically the three camera output signals, sample by sample, with a 3 × 3 transformation matrix. The effect of this transformation is to convert the original camera signals derived from taking sensitivities without negative lobes to the camera signals that would have resulted from proper taking sensitivities with negative lobes.

All this was known to Goldmark and his colleagues. They specified the display primaries and then used the Hardy–Wurzburg analysis to derive the corresponding camera taking sensitivities, including negative lobes. Goldmark understood the need for matrixing to obtain negative camera sensitivities, but he noted that the CBS system had no mechanism for introducing negative values.

Therefore, it is fair to say that by the time Goldmark started his work, the theoretical basis for color TV was fully understood. It was the equipment—the cameras, the transmission apparatus, and the display devices—that was missing.

In the CIE system, in effect, three monochrome cameras and three monochrome displays are used to analyze and then synthesize additive color. The analysis can be successive, with the three-color records taken sequentially one after the other by a single camera equipped with rotating filters, or simultaneous, by means of three cameras and beam-splitter optics. Similarly, synthesis can be successive, with a single display rapidly projecting three images sequentially one after the other through color filters at a high enough rate to avoid flicker, or simultaneous, with the three displays projecting filtered images together in register. Successive methods are simpler to start with because they can make use of black-and-white cameras and displays, which was undoubtedly a primary motivation for Goldmark's approach. (There were no color picture tubes at the time.) Sequential displays work because of the persistence of the human visual system, which integrates images presented at high enough rates. However, the higher the brightness, the higher the repetition rate needed to avoid flicker, a factor that placed strong limits on the performance of the Goldmark system.

It is interesting to note how color television followed color cinematography, which, in the first third of the twentieth century, experimented with various combinations of successive and simultaneous additive systems before the subtractive photographic systems we are familiar with today became predominant [2]. Kinemacolor, which was launched in 1909 and was the first commercially successful process for color motion pictures, was a successive additive system that used rotating color filters or shutters for both capturing and projecting scenes. Because of the practical difficulties of a successive three-color system, Kinemacolor was a two-color system, using orange-red and blue-green filters (roughly aligned with the I axis in the NTSC system).

Because a two-color additive TV system with orange and blue green filters had been demonstrated in July 1939 ([3] in

---

[1] A readily accessible source for abridged versions of the original papers by Young, Palmer, Maxwell, and Ives is [1].

[2] Ives may not have been the first to make this observation.

the Goldmark paper), Goldmark apparently felt the need to point out the limitations of two-color systems. While they are limited in the range of color stimuli they can produce, perceptual effects result in a much wider range of color appearance, as shown stunningly in Land's experiments in the late 1950's [3].[3]

Even with these two-color additive processes, however, the defects of sequential systems were obvious—color fringing (time parallax) and eyestrain from operation near the flicker-fusion frequency. Nevertheless, the development of additive systems of this kind for color motion pictures overlapped the demonstrations of the first primitive color TV systems by Baird in 1927 and Bell Laboratories in 1928 and was pursued for some time afterwards. The former used sequential addition of colors while the latter used simultaneous addition.

III. THE GOLDMARK CLASSIC PAPER

Ideally, a color TV system would use a unitary color camera that mimicked the human visual system, generating three component images that measured three different spectral distributions of light from the original scene, as required to generate accurate tristimulus values. The system would use a unitary color display that mixed three images illuminated by three primary lights. The transmission system connecting the two would convey the information in the three component images, perhaps transformed to obtain some transmission advantage. Goldmark's problem was that none of these three elements existed in 1940, when the work was begun. An additional constraint was that the signal was to be transmitted in a 6-MHz analog channel, such as the one used for the existing monochrome system. His main contribution, which was accomplished in less than two years, was to use existing technology and components to provide an approximation to the functionality of such an ideal system.

Goldmark was familiar with the earlier work in the field, as evidenced by the extensive bibliography appended to the paper. Yet in the paper no consideration was given to justifying the particular system choices that were made or comparing them with alternative methods. He assumed that a trichromatic field-sequential scheme was to be used, and that the three color components to be transmitted would be red, green, and blue.[4]

The only optimization done was to select appropriate scanning frequencies to trade off flicker, spatial resolution, and motion rendition. In the 1942 system, a vertical scanning frequency of 120 Hz was used, along with a horizontal frequency of 18 900 Hz for 315 scan lines per frame and 20 complete frames/s, as compared to 525 lines and 30 frames/s for a standard monochrome image. (There is some confusion in the paper when the author talks of a frame frequency of 50 Hz and both 343 and 375 lines.) This produced a vertical resolution equal to 60% of monochrome and a horizontal resolution of 83% of monochrome. In the version of the system adopted by the FCC in 1950, the frame rate was raised to 24 Hz to reduce flicker at a given screen brightness and to improve motion rendition. A vertical frequency of 144 Hz and a horizontal frequency of 29 160 Hz gave 405 scan lines (the same as then used in Britain) for 77% of monochrome vertical resolution and 54% of horizontal resolution. These unequal numbers seem to have been used because of the hoped-for effectiveness of a horizontal sharpening method developed at CBS [5]. The use of 24 frames/s made transmission from film particularly convenient.

As can be seen from these numbers, a field-sequential system using red, green, and blue frames of equal resolution presents a tradeoff problem that does not really have a satisfactory solution in a 6-MHz channel. The spatial and temporal resolution of the color system had to be substantially lower than that of the existing monochrome system. It is true that color adds apparent resolution in most cases, but it is easy to find subject matter where the lower resolution is easily perceptible. In addition, the screen brightness must be much lower than that of a monochrome system to avoid flicker. Because of this, the inescapable reduction of brightness due to using color filters with a white picture tube, although very large, is not so important. On the other hand, it should be noted that the system did not require a color CRT, it had acceptable color rendition and precise registration of the three component images, and it required no color adjustments at the receiver—features that were not met by the NTSC system for many years after its introduction. Indeed, color sets today still have a number of color controls, although they need adjustment less and less.

Although Goldmark does not specifically mention it, another important system decision was to make no use of storage at either transmitter or receiver, admittedly neither cheap nor easy at the time. As a result of the absence of storage at the transmitter, the three color signals are not simultaneously available to permit matrixing as is required for more accurate color rendition. In addition, color fringing of objects moving rapidly across the field of view is inevitable, since the three color components, being recorded at different times, appear at different places on the screen. The absence of storage at the receiver requires that the three signals transmitted must be of equal resolution, which is wasteful of bandwidth. If the three color components are transmitted at appropriate resolution, experiments show that the overall resolution must be reduced only about 10% (20% on an area basis) to provide fully adequate chromatic resolution, with no reduction in frame rate [6]. Finally, a disturbing effect then called "color breakup" occurred when the eye moved rapidly through the frame.

[3] During the period when the Technicolor Motion Picture Corporation had the only practical three-color motion-picture system, it refused to make cowboy movies. Therefore, those were all shot in two-color systems such as Cinecolor, a subtractive process that used two emulsions, one on each side of the film base. It seems evident, in retrospect, that the apparent color gamut of such systems must have been much higher than a simple-minded application of CIE analysis would indicate. Technichrome, a two-color Technicolor process, was used for filming the 1948 Olympic Games in London.

[4] An excellent short paper by D. G. Fink discussing the advantages and disadvantages of different approaches appeared in 1951 [4].

The paper does go into great detail about the design of the equipment for transmitter and receiver, including studio lighting, the color wheel and its synchronization, camera tubes and CRT's having color characteristics as required, video amplifiers, and timing pulse amplifiers. A "mixer amplifier" was used to permit independent gain and offset controls of the sequential red, green, and blue video signals. Complete circuit diagrams were given for the more important sections. This kind of detail was very common at the time. Recall that the only components available to the circuit designer were vacuum tubes, resistors, capacitors, inductors, and transformers; printed circuit boards were in the future. It was much harder to get even rather simple analog circuits to work properly in the 1940's than it is to get much more complicated digital circuits to work properly today. Project supervisors were heavily involved in minute details of design and construction and had much less time to look at their projects from the systems standpoint.

Goldmark and his colleagues devote a large part of their paper to color. They start at the receiver and then work backward, deriving the transmitter characteristics for faithfully reproducing tristimulus values, which is recognized now not to be a requirement for pleasing color reproduction [7]. Goldmark describes the performance of the receiver as based on the known theory of color vision, while the transmitter characteristics are guided by the desirability of reproducing all colors found in nature. The reverse describes more accurately the important design considerations in a color system: the receiver primaries determine the range of color stimuli that can be reproduced, and the corresponding transmitter spectral sensitivities determine the color matching performance.

The first thing the paper describes is adjusting phosphor mixtures and qualifying black-and-white tubes for use in the receiver so that the receiver, in combination with Wratten color filters, produces the desired white (in this case daylight), which was also used to illuminate scenes at the transmitter. The filters determine the reproduction gamut (and brightness) of the receiver; no reasons were given for the selection of Wratten No. 26, 47, and 58 filters, which are standard color-separation filters used in graphic arts.

The receiver phosphor-filter combinations determine the receiver primaries, which in turn determine the spectral sensitivities required at the transmitter for visual matching. The derivation of the transmitter spectral curves cites Hardy and Wurzburg [8], and although the result reported on p. 185 in this issue is correct, it does not follow from the accompanying text describing how it was obtained. The analysis was extended to take into account "hangover" (residual signals from one color field to the next) on an orthicon tube used at the transmitter. The effect of hangover was to contaminate the color filters so that the red filter appeared to have a leak in the green, for example, which reduces the saturation.

The calculated transmitter color curves, as expected, had negative values at some frequencies. The CBS system, without storage or matrixing, was unable to simulate these curves. The system had a "color mixer," but it was used to mix each individual color signal with various blanking and synchronizing pulses, not with other color signals. The system assumed the same white at the transmitter as at the receiver, and it used the color mixer to adjust camera signals and compensate for scene illuminant changes.

The need for negative values could have been avoided, at the cost of desaturated colors, by using the Ives–Abney–Yule compromise [9], which assumes supersaturated, imaginary versions of the primaries actually used in the system, leading to nonnegative spectral sensitivities for the camera. As it was, Goldmark believed that if negative sensitivities could not be obtained, the best compromise was to use the positive portions of the sensitivity curves [10].

The overall tone reproduction characteristic (gamma) of the system is nonlinear due to the receiver CRT, but there is no mention of gamma correction. Part 2 [10] observed that a gamma higher than one introduces brightness errors but makes the colors appear more vivid. The NTSC system made the same observation and settled on partial gamma correction to compensate for the dim viewing surround, which tends to make colors appear desaturated.

In the end, the authors do not say that their colorimetric analysis was actually used in selecting filters for use at the transmitter, which also used a rotating filter wheel. The filters at the transmitter were almost identical to those at the receiver, reminding one of Maxwell's demonstration of three-color reproduction 80 years earlier. One is left with the strong impression that the colorimetric analysis Goldmark and his colleagues did was used for evaluating system performance but had little or no influence on the system design.

IV. COLOR TV AFTER GOLDMARK'S CLASSIC PAPER

Development of color television systems was suspended during World War II, but many continued thinking about it during that period. When work resumed after the war, researchers (and their companies) were divided into two camps. One side favored the field-sequential system on the grounds that it worked rather well, required only a rather simple receiver, and did not need a color picture tube. Of course, it could use a color tube when (and if) a suitable one were developed. The other side, much more numerous, favored a simultaneous system, expecting that the tube problem would eventually be solved. A simultaneous camera could be built using three monochrome cameras, each sensitive to a different color but having identical fields of view, optically equivalent to those already being used in color cinematography [11]. Lacking a color tube, the display would need three CRT's whose primary color images were either projected in register or viewed directly through beam splitters. Both camera and display would have the problem of accurately registering three scanned images, a problem completely absent from the field-sequential system, but, as we now know, quite solvable.

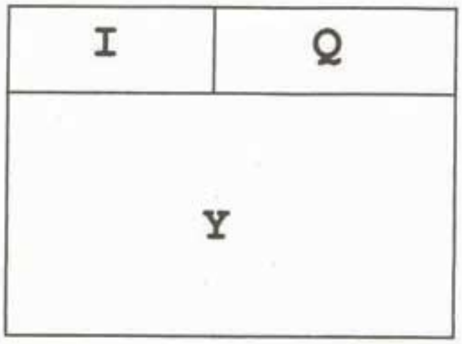

Fig. 1. With hindsight (and storage at transmitter and receiver) it is possible to configure a transmission format that would have been backward-compatible with existing monochrome receivers in 1940 and could have utilized a rotating-filter display. As shown, a portion of the height of the frame would be used for the NTSC I and Q chrominance signals, with the major portion of the screen displaying the luminance signal. This would have given a good picture with wider aspect ratio. The portion of the screen devoted to chrominance would be covered over for color transmissions viewed on monochrome receivers and uncovered for monochrome transmissions on all receivers. With a color CRT, the picture quality would probably have been better than that of NTSC.

Initially, both approaches required three times the bandwidth of a monochrome system of equal resolution. CBS "solved" the bandwidth problem by reducing both resolution and frame rate, a solution that guaranteed noncompatibility with the existing 10 million monochrome receivers. Although it was recognized that, in the long run, compatibility was unnecessary, it was thought by many that it would make it easier to introduce the color system, as the color signals could be received on existing sets.[5] In any event, compatibility could not be achieved in the field-sequential scheme except by using storage at the receiver, a technique that would have greatly raised the cost.[6]

In the simultaneous system, the green signal, if it used standard scanning frequencies, could serve monochrome receivers, while the red and blue signals could be transmitted in additional channels or on subcarriers within the existing channel [12]. Color receivers would use all three channels, unquestionably producing excellent pictures—a not surprising result since the bandwidth was three times as great. The two obstacles to the simultaneous approach thus became the color picture tube and bandwidth reduction. Both problems were eventually solved, but not before the FCC approved the CBS system in 1950. CBS made receivers available and broadcast color programs on a regular basis, but the system never gained acceptance, dying out when the FCC reversed its decision only three years later.

There is a copious patent literature on the subject of color picture tubes, the history of which is summarized by Herold [13]. The first successful demonstration, by RCA in 1950, used the shadow-mask principle. In this method, three parallel electron guns mounted in a tight bundle in the neck of the CRT are made to converge their focused and deflected beams on a perforated sheet mounted a short distance from a glass plate on which are printed hundreds of thousands of phosphor triads. These triads, each having a red, a green, and a blue dot, are arranged in close correspondence to the pattern of small holes in the perforated sheet. If the geometry is right, each gun "sees" only one color of dots, and the problem is solved. This scheme requires such a degree of accuracy in construction and electron optics, maintained over such a wide temperature range, that many thought it would never work.[7]

However, it did work and is still used, particularly for high-resolution applications. Sony's Trinitron uses a similar principle but is easier to manufacture and considerably brighter, particularly for lower-resolution applications. The successful development of such a complicated device is an object lesson in what can be done when highly motivated technologists are supported by a management that has a big stake in the outcome and adequate resources are provided.

The transmission problem was of a different kind in that ideas, rather than technology, played the main role. It had long been known that the red, green, and blue component images did not require the same resolution, although the phenomenon was first described for use in television only in 1940 [14], [15]. Without receiver storage, only the horizontal resolution could be varied. On this basis, the color signal could be transmitted in two standard channels (12 MHz) rather than three (18 MHz), but by the time this was accomplished it was realized that spectrum was in such short supply relative to demand that color would have to fit into one channel. In addition, the issue of compatibility was pressed more and more by opponents of field-sequential, perhaps mistakenly.

Two separate ideas made possible transmission in a single channel as well as compatibility with the existing monochrome standard. Several other ideas improved overall quality. The first essential idea was the "mixed highs" principle [16], and the second was a scheme for transmitting

---

[5] Some observers now think that this was probably not true. When color was eventually introduced, it took ten years to reach 1% market penetration. Color nearly failed in spite of massive support by RCA; the company's exposure was $3–4 billion in today's dollars. Part of the reason may well have been that much of the incentive to buy color receivers, which were considerably more expensive, was lost because the new color programs could be viewed on existing sets. That servicing existing receivers by simulcasting is practical was demonstrated in the United Kingdom and France when 625-line Phase Alteration by Line (PAL) was introduced in 1965. At least some service to old receivers was maintained until 1985.

[6] If the red and blue signals were reduced in resolution vertically as well as horizontally, the green signal would have had to be reduced in resolution only marginally to make space for the other two signals in the same band. All three signals would have had to be interpolated to the same sampling density to be usable with a field-sequential display, and this would admittedly have required expensive frame storage at the receiver. A possible transmission format for a compatible field-sequential system using NTSC color components in shown in Fig. 1. If the signals originated in a simultaneous camera, then the compatible system shown, using a color wheel at first and a color CRT when developed, would have produced results very nearly as good as the eventual NTSC system, entirely free of cross color and cross luminance inevitable with the NTSC format. The loss of vertical resolution in this method would not have been any greater than the loss of horizontal resolution actually experienced in almost all NTSC color receivers, even today.

[7] Including one of the current authors (Schreiber).

the chromatic information within the same 4.2-MHz band used for video in the existing system.

Mixed highs follow from the fact that the three-color components do not require the same bandwidth. Instead of transmitting each signal in a separate channel of appropriate bandwidth, the low-frequency color components are sent in three narrowband channels and the high-frequency portions are added together and sent in a fourth channel. (Low-pass green plus mixed highs can serve for monochrome receivers, although not perfectly.) In a further step, the original RGB components are subject to a 3×3 matrix transformation, giving a wideband (4.2-MHz) luminance signal and two narrowband chrominance signals. This improves compatibility since the luminance signal is ideal for serving black-and-white receivers, and the narrowband chrominance signals can be transmitted within the luminance band with low visibility.[8] The final problem, that of transmitting the two chrominance signals on a subcarrier within the luminance band while still making it possible to separate the signals at the receiver, had actually been solved much earlier by Gray at Bell Laboratories and forgotten [21]. In Gray's scheme, two chrominance signals are modulated in quadrature on a subcarrier whose frequency is an odd multiple of half the horizontal scanning frequency—about 3.58 MHz. This causes the spectral components of the color information to lie in between those of the luminance signal, permitting them to be separated, in principle, by coherent demodulation. Many were skeptical of this idea as well, but it worked well enough to make what came to be called the "NTSC system," adopted by the FCC in 1953, an eventual market success.[9]

V. EPILOGUE

Goldmark deserves a good deal of credit for producing the first workable color TV system of reasonably good quality. Highly inventive and with the ability to carry projects through to successful completion, he made many useful contributions to electronic technology. Unfortunately, like many others, he became a prisoner of his early ideas and was unwilling to look carefully at either the science or technology being developed elsewhere when it conflicted with his own approach. In his 1951 paper, defiantly entitled "Color Television—USA Standard," he did look at alternatives to his filter wheel, giving short shrift to three-tube projection systems that proved to be so successful. Although it did not deserve his valiant defense, the color wheel was not the main drawback of his system. The real problem was that the system ignored then-current knowledge of color perception and, as a result, was wasteful of bandwidth. As far as one can judge from his papers, he did not attempt to deal with this problem at all.

As for NTSC color, it did seem to solve two problems of the field-sequential system—preserving the spatial and temporal resolution of the monochrome system within the 6-MHz channel and permitting existing TV sets to receive color broadcasts in black and white. (It also depended on the development of a practical color CRT.) A significant price was paid for compatibility, in that the horizontal resolution was reduced more than 30% and cross effects were introduced that reduced image quality. A noncompatible system could have retained, or even increased, the resolution and avoided cross effects entirely. There are now more than 200 million TV receivers in the United States, and all of them operate at reduced quality because of the perceived need for compatibility with the very much smaller number of receivers—about 10 million—in existence in 1953. This is particularly ironic in that compatibility is thought by some to have slowed the penetration of color rather than advanced it. Eventually, of course, color became pervasive. It may be fruitful to ponder the great commercial success that NTSC color has had, in spite of its many technical shortcomings.

The country is now in the first stages of an agonizing and very expensive change over to digital TV (DTV) transmission. One lesson that should have been learned from the switch to color is that it is possible to design a much better system without the stifling constraint of backward compatibility with existing receivers.[10]

What apparently was not learned is that the performance of any radically new system must be very carefully examined in all the applications in which it will be used. Before its adoption, NTSC was adequately tested for its only application, terrestrial broadcasting. In the case of DTV, the field testing was totally inadequate, even for terrestrial transmission, and entirely ignored for satellite transmission, although in the latter case the problem is not reliability, but getting the satellite and terrestrial systems to exchange signals appropriately. As far as cable systems are concerned (nearly two-thirds of U.S. homes have cable) all aspects of the digital system were left entirely to the cable people, euphemistically referred to as "the market," again

---

[8] The principle that a color may be represented, not only by a set of three tristimulus values with respect to the red, blue, and green display primaries, but by any linear transformation of these values, was inherent in the 1931 CIE formulation of colorimetry [17] and was also described in detail in [18]. With appropriate (imaginary) primary lights, one of these values is luminance and the other two are chrominance. The use of this formulation for television is generally credited to Loughren [19], who is said to have gotten the idea from a plate in [20].

[9] Actually, the ability to separate luminance and chrominance exactly in Gray's method exists only for still images, and the supposed invisibility of the color subcarrier on existing monochrome receivers depends on the eye averaging the light over exactly one frame. In fact, the subcarrier was clearly visible in areas with highly saturated colors on old receivers that had the prescribed 4.2-MHz video bandwidth. For many years, most color sets used a reduced luminance bandwidth of about 2.5 MHz to avoid this problem. It was only much later, with the development of three-dimensional video spectrum analysis, that the separation problem was fully understood [22]. With this understanding, three-dimensional ("comb") filters can be used to eliminate cross effects. We do not know of any subjective tests that demonstrate that comb filters always improve subjective quality. The real choice is either to tolerate cross color and cross luminance or to reduce the spatiotemporal luminance bandwidth, thus blurring the picture. The NTSC system probably would have been better designed if the luminance bandwidth had been reduced so as to avoid the spectral overlap.

[10] Many in the TV industry did not learn this lesson but insisted for years that high-definition television (HDTV) had to be compatible with existing NTSC receivers. It was the FCC, not the industry, that made the correct decision that the new digital system was to be fully independent of NTSC, and that existing receivers were to be serviced by simulcasting, as was done in Britain and France when PAL was introduced in 1965. Without this decision, digital broadcasting would have been impossible.

without concern for interoperability with terrestrial and satellite systems. Cable may well use a different modulation system from that of terrestrial, raising costs for consumers. The growing use of computers in American homes also received little consideration in the FCC's decision [23].

Finally, we have yet to come to terms with the fact that ostensibly technical issues, such as color TV and DTV standards, are not made in a vacuum but are the subject of intense pressure from stakeholders in the industry. There is no consensus as to how decisions that serve the public interest should or can be made in the face of such pressure. In the color decisions of 1950 and 1953, the FCC, in accordance with its legal mandate, acted in what it believed was the public interest. Looking back nearly 50 years later, their decisions seem quite reasonable. In the DTV case, under the influence of the current popularity of deregulation, many important matters were "left to the market," as a result of which we now have a chaotic situation in which there is a real chance that DTV will fail. This bodes ill for decision-making in our democratic society, which is faced with many problems that are much more contentious than television standards.

APPENDIX I
BIBLIOGRAPHY

1) A. Cornwell-Clyne, *Color Cinematography*. London: Chapman and Hall, 1951.
2) *Proc. IRE*, "Color television issue," vol. 39, no. 10, Oct. 1951.
3) *Proc. IRE*, "2nd color television issue," vol. 42, no. 1, Jan. 1954.
4) W. T. Wintringham, "Color television and colorimetry," *Proc. IRE*, vol. 39, no. 10, pp. 1135–1172, Oct. 1951.
5) P. C. Goldmark, J. W. Christensen, and J. J. Reeves, "Color television—USA standard," *Proc. IRE*, vol. 39, no. 10, pp. 1288–1313, Oct. 1951.

REFERENCES

[1] D. L. MacAdam, *Sources of Color Science*. Cambridge: MIT Press, 1970.
[2] A. Cornwell-Clyne, *Color Cinematography*. London: Chapman and Hall, 1951.
[3] E. H. Land, "Experiments in color vision," *Sci. Amer.*, vol. 201, p. 84, May 1959.
[4] D. G. Fink, "Alternative approaches to color television," *Proc. IRE*, vol. 39, no. 10, pp. 1124–1134, Oct. 1951.
[5] P. C. Goldmark and J. M. Hollywood, "A new technique for improving the sharpness of television pictures," *Proc. IRE*, vol. 39, no. 10, pp. 1314–1322, Oct. 1951.
[6] U. F. Gronemann, "Coding color pictures," Ph.D. thesis, Dep. Elec. Eng., Massachustetts Inst. Technol., Cambridge, 1964.
[7] C. J. Bartleson, "Color perception and color television," *SMPTE J.*, vol. 77, no. 1, pp. 1–12, Jan. 1968.
[8] A. C. Hardy and F. L. Wurzburg, Jr., "The theory of three color reproduction," *J. Opt. Soc. Am.*, vol. 27, no. 7, pp. 227–240, July 1937.
[9] D. L. MacAdam, "Photographic aspects of the theory of three-color reproduction," *J. Opt. Soc. Am.*, vol. 28, pp. 399–418, Nov. 1938.
[10] P. C. Goldmark, E. R. Piore, J. M. Hollywood, T. H. Chambers, and J. J. Reeves, "Color television—Part II," *Proc. IRE*, vol. 31, no. 9, pp. 465–478, Sept. 1943.
[11] A. Cornwell-Clyne, *Color Cinematography*. London: Chapman and Hall, 1951.
[12] R. B. Dome, "Spectrum utilization in color television," *Proc. IRE*, vol. 39, no. 10, pp. 1323–1331, Oct. 1951.
[13] E. W. Herold, "Methods suitable for color television kinescopes," *Proc. IRE*, vol. 39, no. 10, pp. 1177–1185, Oct. 1951.
[14] M. W. Baldwin, Jr., "Subjective sharpness of simulated television images," *Proc. IRE*, vol. 28, no. 10, pp. 548, Oct. 1940.
[15] ——, "Subjective sharpness of additive color pictures," *Proc. IRE*, vol. 39, no. 10, pp. 1173–1176, Oct. 1951.
[16] A. V. Bedford, "Mixed highs in color television," *Proc. IRE*, vol. 38, no. 9, pp. 1003, Sept. 1950.
[17] D. B. Judd, "The 1931 ICI standard observer and coordinate system for colorimetry," *J. Opt. Soc. Am.*, vol. 23, no. 10, pp. 359–374, Oct. 1933.
[18] A. C. Hardy, *Handbook of Colorimetry*. Cambridge, MA: MIT Press, 1936.
[19] R. W. G. Hunt, "Why is black-and-white so important in color?" in *Proc. 1996 IS&T/SID Color Imaging Conference*, Scottsdale, AZ; pp. 54–57, reprinted in *Recent Progress in Color Management and Communications*, R. R. Buckley, Ed. Springfield, MA: IS&T, 1998.
[20] R. M. Evans, *An Introduction to Color*. New York: Wiley, 1948.
[21] F. Gray, "Electro-optical transmission system," U.S. Patent 1 769 920, July 1930.
[22] E. E. Dubois, M. S. Sabri, and J.-Y. Ouellet, "Three-dimensional spectrum and processing of digital NTSC color signals," *SMPTE J.*, vol. 91, no. 4, pp. 372–378, Apr. 1982.
[23] W. F. Schreiber, "The FCC digital television standards decision," *Prometheus*, vol. 16, no. 2, pp. 155–172, June 1998.

William F. Schreiber (Life Fellow, IEEE) received the B.S. and M.S. degrees in electrical engineering from Columbia University, New York, and the Ph.D. degree in applied physics at Harvard University, Cambridge, MA.

He worked at Sylvania from 1947 to 1949 and at Technicolor Corporation, Hollywood, CA, from 1953 to 1959. From 1959 to 1990, he was a Faculty Member at Massachusetts Institute of Technology (MIT), Cambridge, MA, where he is now Professor of Electrical Engineering, *emeritus*. He was Director of the Advanced Television Research Program from 1983 until his retirement in 1990. Since 1948, his major professional interest has been image processing.

Dr. Schreiber is a member of the National Academy of Engineering and has received the Honors Award of TAGA, the David Sarnoff Gold Medal from SMPTE, the Gold Medal of the International Society for Optical Engineering (SPIE), and he is a four-time recipient of the Journal Award of SMPTE.

Robert R. Buckley received the B.Sc.EE. degree from the University of New Brunswick, Canada, the M.A. degree in psychology and physiology from Oxford University, Oxford, U.K., and the SM.EE. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge.

While at MIT he did his graduate work on gamut mapping and color image compression. He is a presently Principal Scientist with the Xerox Architecture and Document Services Technology Center, Webster, NY. Before that he was Manager of Advanced Color Imaging in Corporate Research and Technology. Since joining the Xerox Palo Alto Research Center in 1981, he has worked on color printing systems, color coding and reproduction, image communications, and Internet Fax. He recently compiled and edited the IS&T volume on "Recent Progress in Color Management and Communications."

Dr. Buckley attended Oxford University as a Rhodes Scholar. He is a member of the IS&T, TAGA, ISCC, and he was the IS&T General Chair of the 2nd Color Imaging Conference.

Courtesy of Cliff Benham