THE REVOLUTION AND EVOLUTION FROM DOT SEQUENTIAL TO NTSC

B. D. LOUGHLIN
Hazeltine Corporation

The RCA color television system described in September, 1949, and demonstrated in October, 1949, had a number of innovations and interesting characteristics. But, actually, it was a premature system, requiring further development before it could provide the quality color broadcast service provided by the NTSC color signal. This paper will attempt to trace the evolution and revolution from dot sequential (the 1949 RCA system) to the NTSC color signal.

1.0 THE 1949 RCA SYSTEM

To understand both the revolutionary and evolutionary aspects, we must first understand the details of that 1949 system (1). The basic features of the 1949 RCA system were dot-sequential (i.e., rapid sequential) interface sampling with mixed highs (2). The mixed-highs signal was produced by adding the red, green, and blue signals above 2 MHz in equal amounts. Then, this mixed-highs signal was shunted around the dot-sequential sampling system. The R, G, B signals applied to the samplers were individually limited to a 2 MHz bandwidth and sequentially sampled at a first rate of 3.8 MHz (later 3.6 MHz). The exact rate of sampling was selected to be an odd harmonic of 1/2 line frequency, which provides interlaced sampling. Color sync information was provided by "wiggling" the trailing edge of horizontal sync, corresponding to the interleaved position for color on every other scanning line. The color sampling function was described as narrow angle and specified as having a duty factor of sampling of less than 15% (3).

The receivers demonstrated for the system used sequential narrow-angle sampling to recover R, G, and B information. Again, this was described as having less than 15% duty factor in the sampling. The sampled signals (pulses) were directly applied, using wide bandwidth circuits, to 3 color kinescopes which were viewed through dichroic mirrors to provide a composite color image. Thus, these color images were literally "full of dots" whose annoying effect was reduced by the dot-interleaving relation of color sampling; however, a quite visible crawling pattern remained. Color sync for the receiver was provided by a ringing circuit, excited or triggered by the "wiggled" trailing edge of horizontal sync. Apparently the "full of dots" pictures were what RCA originally intended because the early descriptions say that when tuning in a monochrome transmission, pulses of equal magnitude are delivered to the R, G, and B kinescopes (4).

An alternative receiver arrangement, using shunted mixed highs around the receiver samplers, was described, but there is no evidence it was ever demonstrated.

This 1949 RCA system (when cameras were correctly registered) had the potential for good compatibility between the transmitted signal and existing black and white receivers. However, the color receivers suffered from poor color sync reliability, with resulting poor reliability of color rendition, and produced color images full of crawling dots and other "shimmering" spurious patterns in regions of high detail. In addition, when these receivers were tuned to a monochrome transmission, the dots and shimmering became random in interleaving, and, therefore, much higher in visibility and gave an unsatisfactory monochrome picture. Thus, this 1949 system, while potentially compatible from color signal to monochrome receiver, was really not reverse compatible, i.e., not compatible from monochrome signal to color system receiver. Therefore, in spite of the tremendous innovative efforts put in by RCA's staff, its 1949 color system was premature.

Early in 1950, RCA remedied some of the above-mentioned deficiencies, but the major revolution came with the introduction of the shunted-monochrome and constant-luminance concepts from Hazeltine.

2.0 THE EARLY CHANGES BY RCA

Color sync by wiggling the trailing edge of horizontal sync was proven unreliable early in the field tests/demonstrations. This was replaced by a burst of color sampling frequency during the horizontal back porch. This burst signal was put on a pedestal, so that the combined pedestal and color sampling frequency burst extended from black level to sync tips. This provided a reliable color sync signal which, except for pedestal, is still used today.

Along with color sync reliability came the possibility of reliable indication of no color, i.e., monochrome transmission. This gave the potential of color killing, i.e., disabling of color sampling, upon receiving a monochrome signal. RCA developed a color killer arrangement for their dot-sequential receiver, but this had practical problems. When a narrow angle sample is disabled, large changes in gain through the sampler can occur, making the color and monochrome reception be quite different modes for the receiver. Furthermore, in the preferred form of receiver (the shunted mixed highs form) the lower frequencies, below 2 MHz, came through the "turned off" samplers, but the higher frequencies, above 2 MHz, came through the shunted high channel. Thus, the color killing arrangement additionally faced the problem of matching the two channels in two modes in order to make a flat effective channel, and both good color and monochrome images.

Therefore, while the inclusion of the color sync burst was followed by reliable potential for color killing, the function of color killing of a dot-sequential color receiver was really neither simple nor commercially practical.
3.0 THE REVOLUTIONARY CONCEPT OF BAND-SHARED SIMULTANEOUS

In the spring of 1949, Hazeltine engineers started to give serious thought to compatible color television. Our laboratory efforts were greatly assisted in the early summer of 1949 by RCA providing information on how to build simultaneous color monitors (3 color kinescopes and dichroic mirrors) and color flying spot scanners to obtain three simultaneous color signals. When we visited RCA, Princeton, we knew something secretive was afoot, but information on their new color television system was very carefully kept from us.

During the summer, we, at Hazeltine, were considering and making preliminary tests on simultaneous band-shared arrangements, much like the system later proposed by Bob Dome of GE. Thus, when the details of the RCA dot-sequential system were released in the late summer of 1949, we already had developed somewhat different points of view regarding color television system possibilities.

In September, 1949, I analyzed the RCA dot-sequential signal and realized it could also be considered as being a wideband monochrome signal, plus reduced bandwidth color-difference modulation upon a band-shared color subcarrier. The lower half of the wideband monochrome component (0 to 2 MHz) came via the average transmission of the samplers, and the upper half (2 to 4 MHz) came via the mixed-highs channel. With this concept, the effect of the 3-phase dot-sequential color sampler additionally gave a coloring subcarrier signal (at the sampling frequency) which disappeared on neutral colors (R = G = B). The modulation on the subcarrier produced color-difference signals of R-M, G-M, B-M at 120° points of the subcarrier [where M = 1/3 (R + G + B)]. This modulated subcarrier signal was band-shared because it overlapped the high-frequency monochrome components. Also, the major energy of the modulated subcarrier and the higher frequency monochrome components were interleaved because of the odd harmonic of 1/2 line frequency relation for the sampling.

This new concept led me to realize that both transmitting and receiving equipment could be built more advantageously than for the dot-sequential concept. In the new arrangement, the full monochrome signal is shunted around the sampling. Thus, the samplers became modulators, with the narrow-angle sampling concept having no relevance. The new equipment used two parallel channels, one for the monochrome signal, and one for the coloring signals, or color-difference signals (later to be called the chrominance signal).

Under the new concept, there is nothing sequential about the resulting signal (5). Color-difference signals occur at different phases on the color subcarrier, but the modulation bandwidth is essentially less than 1/2 the color subcarrier frequency, so the color-difference signals, as well as the monochrome signal, are truly simultaneous in nature. There is cross talk between the band-shared channels, and this is low-visibility in nature, due to the odd harmonic of 1/2 line relation averaging in two frames. This effect is the remnant of the former sequential feature.

When originally developed, these shunted-monochrome equipment ideas were thought of merely as alternative ways of processing and developing the dot-sequential signal (6). However, as time went on, we became more conscious of the fact that these equipments really represented a revolutionary change in the dot-sequential system with mixed highs, in which the encoders and decoding equipment were basically different (simultaneous, not sequential) and the only thing the same was the transmitted signal. Even that eventually changed.

Eventually the equipment, its details, and its names, as well as the conceptual annotations, all changed so that such words as: sampling, duty factor, narrow angle, dot sequential, dot interface, disappeared from the lexicon. This revolutionary change in words and equipment was brought about by the shunted-monochrome concept.

George Brown, the engineer in charge of the RCA color effort at that time, says (7), referring to the Hazeltine demonstrations of shunted monochrome in 1950, "I saw immediately that this circuit innovation would give us a precision control over signal adjustments not previously available . . . . . . When we returned home, we immediately changed our transmitting and receiving apparatus to make use of the bypassed brightness circuitry with most beneficial results."

4.0 REVOLUTIONARY DESIGN CHANGES WITH THE SHUNTED-MONOCROME RECEIVER

The shunted-monochrome receiver has design considerations which differ from the dot-sequential receiver in very significant ways.

4.1 Color Killing

In the shunted-monochrome receiver, the coloring channels (producing R-M, G-M, B-M) have zero output upon neutral colors (grays). Thus, to produce a monochrome image upon the reception of monochrome signals, one merely turns off the coloring channel, forcing all color-difference signals to zero. This leaves merely a wideband monochrome signal from the shunted-monochrome channel, and provides a simple color-killing arrangement without need for balancing two split channels as was necessary in the RCA shunted-mixed-highs receiver.

4.2 Saturation Control

The saturation of the colors produced from a shunted-monochrome receiver can be adjusted simply by varying the gain of the coloring (chrominance) channel. This can be done without varying the frequency response to the monochrome component of the signal—which was a problem with the early RCA receivers.

4.3 Color Signal Demodulation

Three samplers were used in dot-sequential receivers. The first shunted-monochrome decoders were also
designed with three demodulators to deliver R-M, G-M, B-M. However, since only two independent variables exist in the coloring subcarrier information, we soon realized that only two demodulators were needed to get complete color information from the subcarrier—such as R-M and B-M. Then, these two signals could be combined in a video frequency matrix to develop the third color difference signal—such as G-M. Most color receivers today still use this arrangement of two demodulators plus a video matrix to get three color difference signals. We referred to this arrangement as the “two-to-three matrix.”

4.4 Delay Equalization

The shunted-monochrome receiver has one wideband channel and a reduced bandwidth color channel. Thus, if uncorrected, the delay through these parallel channels could be different. For optimum transient response, the monochrome channel should be delay-equalized to have the same delay as that of the coloring channel. This was normally done with a short delay line in the monochrome channel. Today, techniques such as CCD delay circuits are available for this purpose.

4.5 Color Trapping

If the monochrome channel has full response at the color subcarrier frequency, the large subcarrier frequency energy on saturated colors can produce desaturation of reproduced colors. (With the tri-color picture tube, moire can also result.) Thus, some attenuation of color subcarrier is desirable in the monochrome channel. This can be by frequency rolloff, or trapping—better sets today use comb filters.

5.0 CHANGE TO CONSTANT-LUMINANCE SYSTEM

Although a shunted-monochrome encoder and decoder produced much superior color pictures, they still had some undesirable spurious effects. Noise was more objectionable in the coloring channel, and areas of fine detail produced very annoying scintillation and shimmer in the color images. This latter effect was due to the band-shared nature of the signal—high frequency monochrome components (formerly called mixed highs) overlapped the color subcarrier region.

The 1949 RCA system had naturally (due to its dot-sequential concept) treated R, G, and B in identical and symmetrical manners. Thus, the color channels had equal gains and equally spaced sampling angles for the R-M, G-M, and B-M signals. Signals and noise in the region of the color subcarrier produced equal magnitude voltage swings in the three-color output signals.

We found experimentally that the visibility of these unwanted spurious effects was far more annoying in green than in red, and even less annoying in blue. This agrees with the order of descending luminance of the respective primaries. This led me to the concept of “constant luminance” (8), in which the coloring channel, of the receiver, is designed to provide cancelling luminance effects between R, G, and B for signals in that coloring channel. This made a major reduction in the annoying visibility of the shimmering patterns in the color image.

In our first experimental tests, we applied opposite polarity random noise to red and green, and adjusted the amplitude of the noise applied to the green for minimum annoyance (9). With this simple test, the balance condition is very dramatic, and the character of the noise changes when mainly chromatic noise (minimum luminance variations) exists, and it is much less objectionable.

Applying the constant-luminance concept generally results in unequal gains in the color difference channels of a receiver—generally being minimum for green, intermediate for red, and higher for blue. The perfection of constant luminance operation obtained is generally limited by items such as: the nonlinearity of the electron guns of the color tube (gamma); difference in phosphor chromaticities from that of the system design, and inequality of phosphor decay time constants. However, early tests showed that 8 to 10 dB more noise could be tolerated under comparable constant-luminance system conditions than with a dot-sequential system. And, again, to quote George Brown, in reference to Hazeltine’s demonstrations of constant luminance, ‘Thus, the final picture was greatly improved over the arrangement we had been using’ (7).

When these asymmetrical decoding changes were made in the receiver decoding, it was found that a specific change is required in the transmitted signal to permit intended colorimetry. That specific change is that the transmitted monochrome signal should be proportioned according to luminance (Y) and not equally weighted [M = 1/3 (R + G + B)].

With complementary changes at decoder and encoder, the constant-luminance system gave intended color pictures with much “quieter” images, because the mutual interference due to band-sharing had been reduced to a tolerable level. And, the combination of shunted monochrome and constant luminance tamed the temperamental and noisy 1949 RCA system into a practical system suitable for quality broadcast use.

This much improved system, together with a special and separate demonstration of the validity of mixed highs, was presented to the FCC by Hazeltine late in April, 1950, and to RCA engineers on the following day. The material was presented not as a revolutionary and different system, but as an improvement to the RCA dot-sequential system. The intent was to convince the FCC that a color TV system following the dot-sequential (vs. field sequential) direction was correct, and that the deficiencies they had been seeing in the RCA implementation had been fixed. In spite of this, the FCC, in October, 1950, rejected the RCA system as impractical, and approved the incompatible field-sequential system as the U.S. color TV standard. This may have resulted from the fact that in spite of Hazeltine’s demonstrations to RCA, and in spite of the recognition by their engineers of the superiority of the shunted-monochrome/constant-luminance system images, RCA submitted a proposed dot-sequential standard, using narrow-angle sampling, to the FCC in the summer of 1950 (3). RCA also published an article on
sampling principles for their system at that time (10).

6.0 PREACHING THE WORD (11)

Late in 1950 or early in 1951, Art Loughren, who was in charge of the color TV work at Hazeltine, decided that a strong educational program should be conducted, to clearly present the technical facts and improvements which appeared to us to make the FCC decision of October, 1950, (approving the field-sequential system) a great blunder. This was the start of a period in which Hazeltine became the teacher of color television principles. This period culminated in publication of a book by the Hazeltine staff, "Principles of Color Television" in 1956. PCTV, as the book affectionately was called, became known and used throughout the world.

During this educational period, both Art Loughren and Charlie Hirsch, second in command of our color work, made many technical presentations to preach the word (11). The second National Television Systems Committee (NTSC) was reorganized to get industry consensus regarding the optimum compatible color television system. Many hours were put in by Hazeltine engineers in NTSC meetings and demonstrations. There were nine panels supporting the main NTSC. Seven Hazeltine engineers were on various panels, and many more involved in demonstrations and in writing technical papers for NTSC.

7.0 COLOR PHASE ALTERNATION

Work on system improvements continued, and Hazeltine and RCA each independently developed a technique known as Color Phase Alternation (12). Under this concept, the color sequence of the color subcarrier is periodically reversed. For example, during one period of time, B-Y may lead R-Y, and during the intervening period, B-Y may lag R-Y. If these periods are arranged so averaging can occur, color hue errors due to phase errors in the system can be averaged out.

At Hazeltine, we investigated both line-rate CPA and field-rate CPA, i.e., reversing after each line or after each field. All experimental work used visual averaging. In field-rate CPA, adjacent lines in space on the picture tube have opposite color errors which clearly average in view of the mixed highs principle. In line-rate CPA, the averaging pattern was more coarse (two lines in space) and had a crawl. No serious thought was given to electrical averaging because we could not conceive that a delay line with a delay equal to one scanning line (for line-rate CPA) could be economically feasible for incorporation in a receiver.

Demonstrations of the averaging effect of CPA, both with regard to large area phase errors and with regard to small area quadrature distortion, were dramatic. A color picture of our demonstration appeared on the cover of Electronics magazine in 1951.

With visual averaging, there appeared little choice between line rate and field rate. Field rate had 30 cycle flicker under large phase errors, but a fine averaging structure under small errors. Line rate had a serious crawling pattern under large phase errors, and a two-line averaging structure under small errors. At Hazeltine, we were still debating line rate versus field rate at the time NTSC was attempting to get agreement on its first color field test signal specification. The Hazeltine engineering memo submitted to Panel 13 of NTSC in 1951, proposing a signal for field test, was written as if either field rate or line rate might be proposed. It mentioned the use of 1/4 line offset frequency for subcarrier with line-rate CPA. Where a specific recommendation for reversing rate was made, the memo was originally written with a blank. An early version of this memo exists with line rate written in, but just before submission to NTSC, we decided in favor of recommending field rate. RCA agreed with the field-rate proposal, and NTSC accepted the proposal, noticing: "It was decided to propose Color Phase Alternation at field rate instead of at line rate because line rate results in annoying line crawl." (From NTSC-P13-162.)

8.0 THE FIRST ROUND OF NTSC COLOR FIELD TESTS

The first NTSC color signal field tested used constant luminance and color phase alternation, and all equipment used the shunted monochrome concept. CPA was at field rate and the color subcarrier was at 3.9 MHz, so CPA was required to average the quadrature distortion transient response due to single sideband transmission on all but the lowest frequencies of color difference signal. The color burst was on a pedestal. R-Y and B-Y were in quadrature, leading on one field and lagging on the other field. The receivers required a field recognizer to properly switch the receiver decoding to agree with the transmitter CPA encoding. This first field test revealed the following.

8.1 Color Sync Pedestal Problem

Early in the testing, horizontal sync instability was observed on a few models of monochrome receivers that were on the market from a prominent manufacturer. Experimental analysis showed this effect was due to response to the pedestal under the color sync burst. At first, concern existed over taking out the pedestal because the burst would intrude into the picture region, where it could light up the image on retrace. Experimental testing with the pedestal removed showed the instability problems were cleared up and indicated no trouble on retrace brightening. Thus, the NTSC signal now exists with no pedestal under the burst.

8.2 High Brightness Displays Versus Field-Rate CPA

Early tests used displays which had clear limitations in peak brightness. RCA began to run tests with small high brightness displays (such as used for projection) and began to be concerned over field-rate CPA. Particularly with the 3.9 MHz subcarrier, all color images had quadrature distortion which had to be averaged out by CPA, and at high brightness and field-rate CPA, these edges flickered at 30 Hz. I do not remember such tests run at line-rate CPA, but, with visual averaging, these bright displays would have had severe crawling patterns.
8.3 Circular Subcarrier

The 30 Hz CPA flicker turned out to be due to less than perfect implementation of constant luminance. Two efforts were made to improve this. One was the use of a so-called circular subcarrier (13). This was a minor modification of the color subcarrier makeup so that the luminance plotted on a subcarrier plane (considering gamma) produced a circle—thus making lumninance-versus-phase be constant in spite of the picture tube gamma. This produced only a small benefit, and the next problem was found to be due to unequal phosphor decays of the red, green, and blue image—which produced a time shift, preventing correct luminance cancellation versus time. NTSC workers found that red and green phosphors could be selected to have medium phosphor decays equal enough to effectively cancel. However, blue phosphors were found to be always very short in decay, much shorter than the desired red and green decay. This limited the benefit of the circular subcarrier in solving the 30 Hz flicker-versus-phasing error with field rate CPA and high brightness displays.

Also, a reduction in color subcarrier to 3.6 MHz was tried, which reduced, but did not eliminate, the flickering quadrate distortion on high brightness displays. And, nobody proposed line-rate CPA with electrical averaging!

9.0 THE OCW (IQ) SYSTEM

Further investigation by RCA and others led to the realization that for medium resolution, the eye is a two-color device. Thus, for coarse detail, three variables are needed to match the eye, for medium detail, only two variables are needed, and for fine detail, only one variable (mixed highs) is needed. This medium resolution axis was determined to be along the orange-to-cyan direction, and was referred to as OCW (Orange-Cyan Wideband) (14)—the wideband being the wider band of the chrominance channel (which has less bandwidth than the monochrome channel).

This dual bandwidth arrangement (1 wider bandwidth, Q narrower bandwidth) provided the potential for eliminating quadrate distortion if the lower frequencies of the color difference signals could be transmitted double sideband on the color subcarrier, and if the wideband axis could be properly selected so as to be suitable for single sideband transmission on the color subcarrier. Laboratory tests to determine the optimum axis for narrow-banding the color subcarrier were considered at Hazeltine, RCA, and I believe, GE. At this time, the circular subcarrier was used, so R-Y and B-Y were not in quadrature, but at 103.6°.

Critical tests showed that with the circular subcarrier, the optimum axis to be narrow-banded was at 40° from B-Y in the direction of R-Y. This angle proved to be not very critical, but a narrow-banding of B-Y was shown not to be the optimum, at least with the color signal material that we had. This is where we became aware of the limitation of Kodachrome (used in our flying spot scanner) and developed masking equipment to attempt to correct for Kodachrome limitations (15).

9.1 Color-Sound Beatnote

At about this time, the color-sound beatnote—a low frequency beatnote produced in nonlinearities like the envelope detector used in receivers—became obvious in a number of black and white receivers on the market. Lowering the frequency to 3.6 MHz did not make this worse, because, although the color subcarrier was further up on the IF characteristic, the beatnote was finer. At almost the last minute, it appeared as though this beatnote problem was going to make the NTSC color signal be marginally compatible. However, a technique of modifying the color subcarrier frequency slightly so the color-sound beatnote interleaved, under zero sound modulation, saved the day. This resulted in a very slight change in field frequency in order to keep the color subcarrier an odd harmonic of 1/2 line, as well as color-sound beatnote another odd harmonic of 1/2 line, and keep the sound frequency still at 4.5 MHz from the picture carrier frequency.

9.2 The Last-Minute Change

Certain NTSC members pushed for returning to R-Y and B-Y in quadrature. All the OCW (IQ) tests were done with a circular subcarrier. A subcommittee studied this question, and in view of the relative "non-criticalness" of the optimum narrow-banded axis, calculated the equivalent angle for the change from circular subcarrier to R-Y and B-Y in quadrature. This very uncritical optimum of 40° in the circular subcarrier converted to 33° in the quadrature case. The number 33° appears to have much greater stature than is proper. Probably ±10° from the narrow-band axis would be difficult to see.

10.0 RECAPITULATION

The premature 1949 RCA dot-sequential color system progressed to NTSC by:

1) use of color burst instead of trailing-edge wiggle for color sync;
2) change from a dot-sequential to a band-shared simultaneous system by modification of transmitting and receiving equipment to incorporate shunted monochrome (and then using the correct lexicon);
3) addition of constant luminance concept to reduce the visibility of band-sharing cross talk;
4) removal of the pedestal from under the color burst;
5) taking a little side adventure into CPA at field rate, which was field-tested and removed;
6) replacing CPA with the IQ or OCW system with three different bandwidths;
7) minimizing color-sound beat note by making it an odd harmonic of 1/2 line frequency;
8) making a final change of circular subcarrier back to R-Y and B-Y in quadrature.

Step 1 came from RCA. Steps 2 and 3 came from Hazeltine. Steps 4 through 7 were all made through NTSC at the suggestion of either RCA or Hazeltine. I don’t remember who originated Step 8.
11.0 POSTSCRIPT

More than a decade after NTSC took CPA out of the NTSC signal, Dr. Bruch of Telefunken came forward with PAL (16), which is CPA at line rate. In 1951, line-rate CPA was considered, and, at least at Hazeltine, so were certain other features now in PAL, like: 1/4 line offset of color subcarrier and the oscillating burst to provide alternate-line sync (17). However, the key of electrical averaging by a delay line, to the best of my recollection, did not receive serious consideration in NTSC, probably due to the perceived economic impracticality (at that time) for commercial receivers. This seems to be a case where the development of a certain critical component could have affected the course of color TV system development (in U.S.A.). Many engineers today consider delay-line PAL to be "more rugged" than NTSC.

REFERENCES

4. Reference 1, p. 516.
7. G. H. Brown, "and part of which I was," Recollections of a Research Engineer, Angus Cumar, pp. 210, 211, 1982.

BIOGRAPHY

B. D. Loughlin's many achievements include: Cooper Union, B.E.E., 1939; E.E., 1945; Professional Achievement Citation, 1969; Gano Dunn Medal; 1970. Stevens Institute of Technology, M.S.E.E., 1946; Hazeltine Corporation, 1939 to date; Vice President, Research 1967–79; Vice President, Technology since 1979; Institute of Electrical and Electronics Engineers, Fellow; Zworykin Television Prize Award; Professional Group on Broadcast and Television Receivers Award; Consumer Electronics Award; Society of Motion Picture and Television Engineers; David Sarnoff Gold Medal Award; Special Commendation Award; National Association of Manufacturers, Modern Pioneer Scroll Award; International Television Symposium Citation; Member, National Academy of Engineering; 109 U.S. Patents assigned to Hazeltine.