High/Low Density Information Storage Using Colors

Harry A. Shamir
ColorCode, UnLimited Corporation, Cambridge, Massachusetts

Technical Abstract

ColorCode UnLtd. Corp. is innovating a method of using shades of colors to store data in a distributed, secure, machine readable, low cost, compressed format. The technique lowers storage cost, and processes all knowl-edge bases: text, data, images, sound, even olfactory and tactile. The information is encoded using special algorithms into an array of superdots (hueDots™). Decoding algorithms retranslate the information into comprehensible form. The databases and information are stored on paper, plastic, film, or photopaper.

The methodology hinges on the accuracy of solid state technology in detecting properties of light. A color calibration technique compensating for distortions due to printer, reader or media and eventual fading of colors, is built in.

ColorCode offers several formats for widespread, multi-media, distributed information storage, including miniature labels, rich in information. The latter mode of usage is a major advance in manufacturing management, especially for aerospace items.

ColorCode's methodology has user-friendly information properties with substantial economic and social benefits, and is technically innovative. Uses are foreseen in:

1. Storage, use and distribution of data directly and physically on parts and assemblies for aircraft, space craft, automotive, and others. Net result is retention of machine manuals on the parts themselves. The physical form would be ColorCode labels printed on one of several types of media.
2. Paperless Publishing and ID cards: information storage in a small area carrying large amounts of information.
3. Long term storage of large quantities of information on a durable medium for archival purposes.

Introduction

Modern data processing systems have recognized the utility of color and color patterns for the human-machine interface. Few users still have black and white screens on the computer, for example, and color displays are used in industry to alert an operator to significant events.

As the computer has evolved, so has image processing. This is the translation of images and sounds into bits of information which can be stored and manipulated within the computer.

ColorCode marries these two technologies to move the use of color beyond the interface of man and machine into image processing and data storage. The CC technique addresses two basic needs: (1) the need for usable information accessible in a very distributed fashion; and (2) useful management of databases, by making them accessible at the points of need, ***with immediacy***.

Implementation of ColorCode technology results in: (1) large amount of information stored upon a small area of any article or part; (2) long term, large volume information archival; and (3) intermediate quantities of information, stored on film media and made available by mass production.

Current technology for information exchange and distribution requires either direct interconnection among computers, or transfer of physical storage media, e.g. disks, paper or film. The current highest density of physical information storage repository systems use ablated microdots (optical means) on metallized plastic media, implemented today in CDs, CD-ROMs and Optical Tapes.

In the Distributed Information industry, ColorCode is convinced a market exists for a storage technology based on color imaging. ColorCode innovates technology enabling access at the points of need, ***with immediacy***. ColorCode believes a large body of uses will be labels on aircraft and aerospace parts and assemblies, followed by automotive parts and assemblies.

An additional industry that could be served is the information archival industry including bibliographic and databases repositories, with **substantial** cost advantages over current technology. An additional implementation field is the paperless publishing industry, deriving tech-nology from the archival application.

ColorCode's plan is to use adhesive backed labels printed with hueArrays for low cost labelling of aerospace parts. Low cost archival of data for serial and random access will use photo-film. Paperless publishing will be based on cards containing information in a highly compressed format.

Background

A major current optical storage technology uses ablated microdots, and is implemented today in the forms of Compact Disks (CDs), CD ROMs, 'laser-cards' and optical tapes. The microdots are formed by ablating or physically altering microspots of material to record binary digits.

The major players in current optical storage technology are SONY, 3M, and others, that offer large disks to 2" diameter CDs, and Drexler Technology that offers Lasercards. Each device or disk size requires its own player, which is a relatively large machine.

We explored the technical possibility of recording data by the microdot technology onto a label format. We found that no such technique currently exists, nor is any envisioned. Therefore this technology is not currently competitive with ColorCodes' vision of using hueCoded information storage cards and on labels to be adhered to parts, assemblies and documents.

We explored the possibility of using small optical disks to mass distribute data for the paperless publishing market. We found that even the smallest disks that carry no more than 2-4 Megabytes, require Readers too large to be easily incorporated in laptop and palmtop computers.

We explored the costs of archiving data as presently served by microdots technology on spools of tape, each containing 1 Terabyte (1 million megabytes) of data. Current technology archival costs are $250,000 for a Writer/Reader, and $10,000 per 1 Terabyte reel (10^{12} Bytes). ColorCode projects that archival technology based on high density color film technology is estimated at half the above costs.

ColorCode is not the first entity to have proposed the concept of storage of information using colors. The coding of resistors using color bands is well known. Photography and films have been suggested and studied to carry information based on their ability to store shades (intensities) of colors.[1-12] However, ColorCode is the first to address the reduction of the distortions, recurrent noise, slow drift, and color fading. Such reductions bring the storage of information using colors from the realm of interesting possibility to practical and commercial reality.

Technical Approach

Colors do not exist outside the eye-brain. What does exist are signals of given wavelength having given signal amplitudes. A collection of these signals is called 'bandpass'. A color filter enables transmission of a specific bandpass, where not all the component wavelengths are transmitted with equal ratios of input to output. Color technologists have worked long and hard to create filters that mimic the filtering characteristics of the average human eye. Their products are available on the marketplace at acceptable price, and are used by ColorCode in conjuction with machine vision and light measuring technologies.

The company's technology focuses on the use of color (including gray) and color intensity levels (i.e. shades) as the key for encoding of data. The information is carried in a hueArray™ of superdots or hueDots™ (a dot of composite color). The information is encoded, and optionally encryted, using special algorithms by a technique called hueCode™. A decoding algorithm with optional deciphering, present in the reader's host computer, translates the information back into a humanly legible form.

Storage of Information on One Superdot

An array of hueCodes is called a 'hueArray'. The constituent colors of each superdot will each be within a wavelength bandpass, and have one of a set of intensities.

Let C = Color, A = a set of wavelengths (colors), I = intensity of the light of a specific wavelength. Clearly, the continuum spectrum of visible light and proximate wavelength ranges can be subdivided into contiguous band-passes of arbitrary width:

UV | | | | | | | | | | | IR

$A_n = Q_1$ S_1 T_1 R_1 Q_2 S_2 T_2 R_2 Q_3 etc.

We shall arbitrarily use only the Q set of wavelengths. The reason is that we choose to leave a wide margin around each wavelength band for prevention of error.

Each wavelength A_n can be recorded in a variety of intensities I_k. The result is a matrix Q_n x I_k:

Color	Intensities
Q_1	I_1 I_2 I_3 .... I_k
Q_2	I_1 I_2 .... I_k
.	.
.	.
.	.
Q_n	I_k

We can incorporate on one superdot **n** different wavelengths, each having one intensity from the available population of **k** intensities. Practically, existing commercial printing techniques limit the choice of wavelength bands. Color cameras exist, able to distinguish the three colors red (R), green (G), and blue (B). Therefore we use RGB where each constituent color can have one intensity level out of k different levels. Modern technology may enable k to be as much as 16 or 32 or 64. Electronics can achieve much more, but we deal with the real world of dyes and pigments, media and illumination, thus understand only too well the restrictions from practicality. However one can hope, thus we will calculate based on k = 64.

Given a number of wavelength bands **n** = 3 and a number of intensities **k** = 64, then the number of available combinations for each superdot is $(64)(64)(64) = k^n = 64^3 = 262,144$. This is in essence a three 'digit' numeric ele-ment of the set of numbers of base 64, equivalent to 18 independent 'bit' values, as will be shown below.

To reduce this somewhat esoteric number to more familiar themes, we shall calculate it in terms of bases: 10 (the familiar 0 through 9), 256 (the number of ASCII codes), 26 (the letters of our alphabet, a through z), and 2 (the binary code):

$$64^3 = 256^s = 26^q = 10^p = 2^r = 262144$$

Calculating, we obtain: $s = 2.25$
$q = 3.829429 = \sim 3.82$
$p = 5.4185399 = \sim 5.41$
$r = 18$

Each superdot can hold 18 bits of information. This translates into

2 1/4 bytes
> 3 ASCII characters
or > 5 numerals per superdot

If more than 3 colors were used then the available storage per superdot would be increased. e.g. for 4 colors at 64 intensities each, the number of available combinations is 16,777,216, and the number of bits per superdot is 24.

Storage Density of hueArrays

The amount of information stored in a hueArray will depend not only on the number of colors but also on the size of the huecode dot (spacial resolution) on the film, the size of the array, the dynamic range (color resolution), the arbitrarily chosen factors of redundancy in the system design, and other factors.

As a simple example, an array of 20×20 huecodes can represent a string of 2164 base 10 numerals (0 through 9), out of an available population of $(262144)^{400} = \sim 2.606 \times 10^{2167}$. In practice, some of the available string of 2164 numerals will be allocated to error checking and other internal data management purposes. We estimate the number of usable numerals will be about 1400.

The superdots in the label on the bottom of Fig. 1 are 500 microns x 500 micron. There are 240 dots in this matrix, representing 740 ASCII characters. This is approximately the same information content as one quarter of a single spaced 8 1/2" by 11" page of text, assuming the 2000 characters per page from widely used word-processing systems.

In contrast, if film based huedots were 3 microns in diameter (or square) with 2 micron spacing, each dot-space would occupy an area of approximately 25 square microns. This would yield a dot density of 25 million dots per square inch. At this density, allowing for 2 1/4 ASCII characters per dot, the information density is 56 megabytes per square inch. *An excess of 28,000 pages of information could be coded and placed in a single square inch.*

Compare this with the information density of a CD ROM (WORM), which on a 4" diameter disk is at best 725 megabytes, yielding 32 megabytes per square inch, which is only *57% of the information density of the level G hueArray,* (see Table 1).

Table 1. Storage Capacity for 1 sq. in. Area

level*	hueDot size (microns)	hueDot spacing (center to center) (microns)	# hueDots (no./sq.in.)	ASCIIs (capacity per sq.in.)	Text Pages (per square inch)
A	450	500	2,500	5,600	2.80
B	230	250	10,000	22,500	11.25
C	90	100	62,500	140,500	70.00
D	45	50	250,000	560,000	280.00
E	22	25	1,000,000	2,250,000	1,125.00
F	8	10	6,250,000	14,050,000	7,000.00
G	3	5	25,000,000	56,000,000	28,000.00

For maximum information density, the dot sizes (which here include spacings) must be minimized, but optimized so the data encoding accuracy and reliability are not compromised. Table 1 depicts the information storage capac-ity of an array of hueCodes that would fit onto a 1" by 1" area at various dot size and spacing levels.

Tools

Storage of information on film at level F occurs using widely available color film recorders and commerical color films. Though not providing the ultimate in dot densities of level G, these machines are acceptable hue-Array printers at level F. The conservative estimate for the center to center distance of hueDots using the claimed resolution of the Agfa Matrix Forte (film recorder) of 2742 rows at 4096 pixels per row, is .009 mm (9 micron). The machine issues 24 bits per pixel. This in theory will meet the requirements of level F information density, and even surpass it. However, the machine does create a 25% overlap of dot edges, decreasable by performing repeated shots at lower light intensities. It remains to be seen if these machines can be 'pushed' to a higher information density by using higher resolution ('slower') film and/or improving the lens system on the camera assembled to the machine, and/or altering the software. One of the goals of ColorCode is to achieve level G information density on transparent films. The maximum achievable information density using reflectance films (photopapers) remains to be checked. ColorCode expects to achieve levels B^+ or even C^- using instant photography films.

Commercial Readers for films are also widely available. Some flat bed paper color scanners have sufficient resolution to serve ColorCode purposes. To act as hueArray readers, these scanners require only ColorCode software.

Prior Art

As mentioned earlier, ColorCode is not the first to have conceptualized the storage of information using colors. However, prior investigators[1-12] did not have the advantage of modern computerized information analysis. Therefore, little if any work was done to separate random and rapid noise from effects of slow drift, and non-linearities in transformations from one medium to another, which are distortions recoverable by calculations. In 'Related Research' below, is listed work performed by ColorCode principals and associates, indicating the advent of powerful techniques able to correct for signal distortions such as listed above. The incorporation of calibration dots in the hueArrays ensures the applicability of such techniques for recovery of information heretofore deemed irretrievable.

Prior researchers also erred in concluding that storage of information using colors have inherently a low signal to noise ratio. Among the errors were the assumptions of linearity inherent in the use of Fourier transforms in the modelling of photographic information storage.[12] Photographic imaging is well known for its non-linearity.

Achievements and Current R&D

ColorCode has demonstrated the general concept of storing information using shades of gray and colors on a variety of media. Practical limitations to the technology are set mainly by the characteristics of commercial Printers, Scanners and media capable of achieving the desired Resolutions.

Current activity includes improving information density using laser-printing, xerography, film recording, and print-on-demand technology in monochrome and color.

In the xerographic and dye transfer implementations, commercial and inexpensive media are used, including paper, PVC, and others. ColorCode is currently active in R&D in these fields.

Another area of interest is adhesives. The search is for contact adhesives that do not lend themselves easily to removal. The purpose is to foil counterfeiters. Other adhesives are sought that also are difficult to remove, except with use of specific removal agents, whose nature would be a closely held trade secret. Removability of the labels by destruction is inherent in the tamper-proofness of the hueLabels. To remove adhesive residue normally requires solvents. Doing this in a factory is commonplace rework procedure. However, the use of solvents is messy. ColorCode is exploring a solventless approach which is still confidential.

Another area of activity is the marriage of the hueLabel with foils. The latter are used in anti-theft devices, activating alarms when passing through special portals. The surface area of the labels carrying these foils can easily be adapted to carry hueArrays of various types.

Future ColorCode R&D activity will include developing Levels F and G information densities.

In the photographic implementation, ColorCode uses media such as color film and photopaper available commercially for human eye consumption to store information, thereby ensuring low media cost. Further experimentation will lead to conclusions regarding what can be expected from improved imaging hardware, used with higher resolution 'slower' (lower ISO number) commmercial color films.

Anticipated Benefits and Markets

Early applications for ColorCode technology are envisioned in Distributed Information by use of levels A through E cards and labels on documents, parts and assemblies, in the aerospace, medical, machine, automotive industries and others. These labels would track and record part quality history in real time during the manufacturing process, and provide *instant access in field* to maintenance, rebuild, refurbishing and recycling instructions.

ColorCode technology has high potential payoffs in 4 major areas:

1. Levels A^- through D information density: Storage, use and distribution of data directly and physically on parts and assemblies for aircraft, spacecraft, automotive and others. Net result is retention of machine manuals on the parts themselves. The physical form would be of ColorCode labels, printed on one of several types of media.
2. Levels F & G: Long term storage of large quantities of information on a durable medium for archival purposes.
3. Level G: Paperless Publishing, in conjunction with a miniature reader, used in a laptop-like computer for multimedia interface with humans.
4. Cards and labels containing information at densities lower than A, created very inexpensively by laser-printers, used in personal Medical Record and ID Cards, automotive registrations, licenses, travel documents, corporate cards, and more.

In all cases the attraction to the industries will be due to lower costs, higher capabilities, and long term survival of the technology.

Aerospace Industry

The aerospace industry using hueLabelled parts will reap the following benefits:

1. *Field access with immediacy*, of data for each and every part and assembly used in the industry, without searching in manuals.
2. Obsolescence of paper based maintenance tomes.
3. Non reliance on long lines of communication, with their concomitant delays, interruptions in service, and warpage in communication results.
4. Traceability of each and every part and assembly in the industry, regardless of its geographical location.
5. Paperless documentation of all manufacturing steps and quality control results for each and every component used in the industry.
6. On site verification of authenticity of parts.

All these factors reduce very costly equipment downtime.

Multimedia Paperless Publishing

To become a reality, multimedia paperless publishing will require portable and lightweight Readers, descendants of today's laptop computers. Two factors presently inhi- bit the birth of such Readers: the lack of commercial low cost thin and lightweight High Definition TV (HDTV) screens, and concomitant minute electronic platforms.

HDTV is expected to become widely available in the U.S. within a few years. The advent of the compact electronic platform is inhibited by the lack of a minute, thin Reader capable of scanning 2 to 4 Megabytes of data per each entry. A small, thin Reader predicates an even thinner information carrier. Film is quite thin. If it can be proven to be an effective high density information carrier, film can potentially become the carrier of choice.

Summary

ColorCode is offering the market a range of products for information storage featuring low physical volume, ease of information retrieval and low cost. Use of general purpose printers and scanners enables affordable multimedia equipment. HueCoding becomes a general purpose tool, providing each type of user an optimal solution to his/her specific problem.

References

1. C. S. McCamy, " On the Information in a Microphoto graph", *Appl. Optics*, **4**, p. 408 (1965).
2. N. Bar-Chaim, A. Seidman, E. Wiener-Avnear, "A Color Memory Mode Based on the Variable Birefringence in PLZT Ferroelectric Ceramics", *Ferroelectrics*, **11**, pp. 385-388 (1976).
3. R. A. Bartolini, H. A. Weakliem, B. F. Williams, "Review and Analysis of Optical Recording Media", *Opt. Eng.*, **15**, pp. 99-105 (1976).

4. F. N. Magee, "Discussion of Electrophotographic Film", *Laser Recording and Information Handling*, A. A. Jamberdino, Ed., Proc. SPIE 200, pp. 16-19 (1979).
5. S. Masdlowski, "High Density Data Storage on Ultraviolet Sensitive Tape", *Optical Storage of Digital Data Technical Digest*, Opt. Soc. Am., p. WA4-1 (1973).
6. R. G. Zech, "Review of Optical Storage Media", *Optical Information Storage*, K. G.Lieb, Ed., Proc. SPIE 177, pp. 56-62 (1979).
7. Y. Makunoki, *et. al.,* U.S. Pat. # 4,278,797 (1981).
8. R. C. Jones, "Information Capacity of Photographic Films", *J. Opt. Soc. Am.*, **51**, pp. 1159-1171 (1961).
9. R. C. Jones, "On the Point and Line Spread Functions of Photographic Images", *J. Opt. Soc. Am.*, **48**, pp. 934-937 (1958).
10. J. H. Altman, H. J. Zweig, "Effect of Spread Function on the Storage of Information on Photographic Emulsions", *Photo. Sci. Eng.*, **7**, pp. 173-177 (1963).
11. R. Shaw, "The Application of Fourier Techniques and Information Theory to the Assessment of Photographic Image Quality", *Photo. Sci. Eng.*, **6**, pp. 281-286 (1962).
12. D. R. Lehmbeck, "Experimental Study of the Information Storing Properties of Extended Range Film", *Photo. Sci. Eng.*, **11**, pp. 270-278 (1967).

Related Work

J. Schwartz, T. T. Wu, U.S. Pat. #4,658,286, filed 3/28/83, "Method and Apparatus for Correcting Distortions in Reproducing Systems".

J. Schwartz, T. T. Wu, U.S. Pat. #4,511,229 filed 6/9/83, "Method and Apparatus for Evaluation of Color Information in Photographic Processes".

J. Schwartz, "Color Equalization", *Journal of Imaging Science & Technology*, July-Aug. 1992.

H. Shamir, "Multi-Color Information Encoding System", U.S. Pat. Pending.

H. Shamir, "Microlabelling System and Process for Making Microlabels", U.S. Pat. #5,118,369.

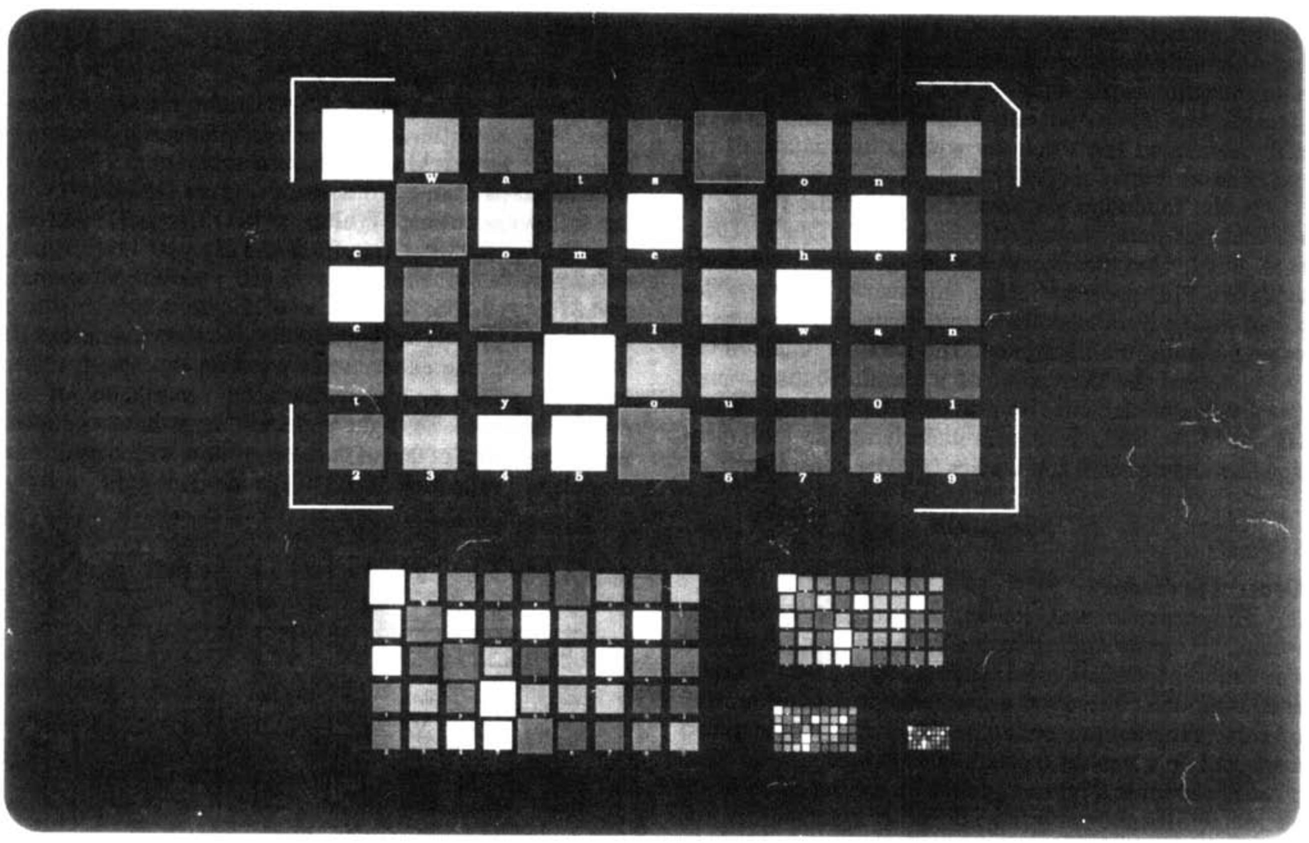

Figure 1. Note: The original of this figure is in multi-color.