INTRODUCTION TO VISION, COLOUR MODELS AND IMAGE COMPRESSION

Slavik D Tabakov

Dept. Medical Engineering and Physics, King's College London, UK, Vice-President IOMP

Abstract— The paper presents an introduction to the human vision and perception of colours. Short examples are given for spatial resolution of a normal eye and its colour sensitivity. Several colour models are briefly presented with emphasis on Hue Saturation Value (HSV) model. The concept of the latter is shown as underpinning the implementation of digital image compression. Elements of the paper can be used as introductory to various educational modules/courses on medical imaging.

Keywords— human vision, colour models, image compression, medical imaging educational resources

TRICHROMATIC COLOUR REPRESENTATION

Colour theory is in the focus of imaging sciences from the time of Leonardo da Vinci. Some milestones include Isaac Newton's Theory of Colours (in Opticks, 1704), Michel Eugène Chevreul's Law of Simultaneous Color Contrast (1839), Young–Helmholtz Trichromatic Colour Vision Theory (1850) [1], and James Clerk Maxwell's Colour triangle (1860). These fundamental scientific works were supplemented by works of philosophers and artists such as Goethe (Theory of Colours, 1810), Schopenhauer's On Vision and Colors, and importantly, Albert Munsell's artistic Color System and Notation (1912) [2].

The complexity of the Colour theory is based on the mix of scientific (light), physiological (vision) and psychophysical (perception) components [3]. Initially Newton's discovery that colours could be created by mixing of colour primaries led to selection of three "pure" or "primary" colours, which were accepted as Red, Yellow and Blue (RYB Colour model). Goethe also adopts these "primary" colours and today many artists and painters consider the RYB as "primary". In parallel Thomas Young suggested that the eye includes three specific colour receptors (1802) and later (1850) Hermann von Helmholtz developed these ideas further to suggest that these three types of photoreceptors (later named cone cells) could be classified as short-preferring (blue), middle-preferring (green), and long-preferring (red), as per their response to various light wavelengths. This established the RGB colour model, used extensively at present. The actual cone cells in human retina were described one century later by Gunnar Svaetichin (1956) [4].

Most of those pioneers accepted that the mix of three primary colours can reproduce all colours in the light spectrum. Any departures from this assumption were explained with the impurity of the primary colorants (dyes) used. Today we know that 3 primary colours can only create a limited number of colours (a gamut). This number is smaller than the human eye can perceive, but is more than sufficient for reproduction of colour in prints, monitors, etc. The trichromatic theory led to the use of the RGB model for television and John Baird created the first colour TV transmission (1928) using Red, Green and Blue emitting phosphors [5]. This established the RGB as the technical "primary" colours.

The fact that active displays (projecting light, as TV monitors) add various proportions of Red, Green and Blue light intensity to form a specific colour from the gamut, led to naming these colours "additive". The addition of all three colours in their maximal intensity will produce almost white light. In contrast, pigments which absorb these specific RGB colours from the spectrum of the light they reflect, were named "subtractive" colours. The subtractive colours of Red, Green and Blue are respectively Cyan, Magenta and Yellow (CMY) – see Figure 3. These are sometimes called "subtractive primaries", or "secondaries", or incorrectly "negatives" (following the photographic white-black analogy). Light reflected from maximal concentration of CMY pigments will be almost black as its RGB components will be absorbed by the pigments. The CMY subtractive colour model is used for colour printing, but black (K) is added as a separate pigment (to avoid use of too much colour pigment ink) – the so called CMYK model.

The need of better colour digital printouts (more colour nuances) required the inclusion of additional subtractive colours (e.g. light cyan, light magenta) forming a six-colour printing (aka CMYKLcLm), while the additive colours remained almost unchanged. Today all colour digital monitors form their colour pixels by three active elements with Red, Green and Blue glass filters over their emitting area, and similarly the photosensors of most digital photo camerasare covered with similar RGB glass filters – the Bayer filter. This filter includes 50% Green and 25% for each Red and Blue to reflect the increased sensitivity of human eye to green light [6] – see Figure 1.

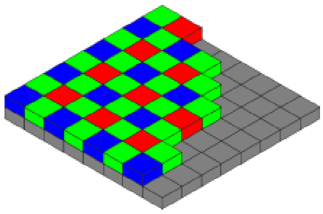


Figure 1. Bayer filter (image cited from [32])

COLOUR SENSITIVITY AND RESOLUTION OF THE HUMAN EYE

The discovery of the light sensitivity of the cone cells and rod cells in the human eye retina (G. Wald & P. Brown at Harvard, and E MacNichol, W Marks & W Dobelle at John Hopkins, 1964) placed the RGB model on more solid scientific background [7,8]. Although the photoreceptors of the retina had been known since the 19 century and the light-sensitive protein rhodopsin had been discovered by Franz Boll in 1876, the more detailed understanding of colour perception needed some 100 years further research. Despite this, the visual perception and the colour processing in the human brain are still not well understood.

What we know currently is that the three types of cone cells in the retina of the human eye have the following response to light [9]:

- Blue sensitive cone cells (aka S cones, for Short) with peak sensitivity around 420-440 nm
- Green sensitive cone cells (aka M cones, for Medium) with peak sensitivity around 534-555 nm
- Red sensitive cone cells (aka L cones, for Long) with peak sensitivity around 564-580 nm

The other type of photosensors in the retina - the rod cells - have sensitivity between Blue and Green (but closer to Green) with peak at about 498 nm – see Figure 2. The rods are very sensitive to light (about 100 times more sensitive than cones) and provide vision at low intensity levels (twilight and night vision) [10].

The rods in the human retina are about 120 million, compared with cones, which are about 6 million in total. The cones are mainly concentrated in the macula of the retina and are larger than rods (cone diameter is c. 0.006 mm, while rod diameter is c. 0.002 mm) [12]. The distribution of the cone cells is roughly accepted as 64% reds, 32% greens and 2% blues (also, their distribution in the retina is not homogeneous) [13]. Due to this reason the maximum sensitivity of the human eye is in the region of the green-yellow colour. As the cones are responsible for perception of the high-resolution images, the representation and transmission of colour is directly related not only to the visual contrast resolution, but also to the spatial resolution.

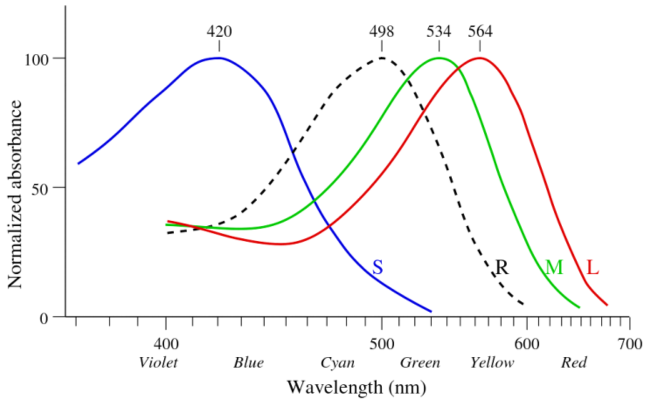


Figure 2. Response of cone cells (S, M, L) and rod cells (R) to light wavelength (image cited from [10], based on [11])

If we take a rough example [14] of reading text from 50 cm distance (book to eye lens) and assume c. 2 cm distance from the eye lens to the retina, the observed text (image) will be minimized 25 times over the retina. Projecting the size of the cone (0.006mm) over the text, will present an object of 0.006x25=0.150 mm (i.e. a pair will be 0.3 mm). This object size corresponds to 3.33 lp/mm (line pairs per mm). For printing, about 170 such objects will be displayed over 1 inch – i.e. 170 dots per inch (dpi). This indicative example gives an estimate of the minimal acceptable spatial resolution (without zoom or optical magnification) of observed printouts or films with normal eye at 50 cm. However the vision acuity depends on the observation distance. Usually the visual acuity is expressed in cycles per degree (cpd), this taking into consideration the viewing angle (and the distance from the object). As every optical system, human eye has its Modulation Transfer Function (MTF), its Contrast Sensitivity, Subjective Quality Factor, etc – for more information on the subject see [15].

The colour sensitivity of the human eye depends on various parameters, most notably the luminance. Usually it is accepted that a normal untrained eye could distinguish between 150 and 250 different grey shades, while the number of distinguishable colours in this case can be from 100,000 to several millions [16]. As medical imaging relies mainly on grey scale images, research has shown that a well trained eye could distinguish around 870 just noticeable differences (JND) of grey [17]. The pixels of contemporary digital medical images use 16 bits, of which usually 12 bits are used to record the image contrast, and the other 4 bits are used for supporting information (for example text or graphs displayed over the image). The 12 bits present 4096 levels of grey (2^{12}) what is more than enough for the human vision. A special Windowing technique is applied to adjust this large number of grey levels to the less-sensitive human eye [18]. Historically the selection of 4096 grey levels has been based on the early CT scanners, where the difference between CT numbers of air and water has been accepted as 1000, while the most absorbent bones are up to 3 times this absorption difference, thus forming a CT number scale from -1000 through 0 (water) to +3000. The practice has shown that 4096 levels of grey are also sufficient for various densitometric measurements (measurement the optical density of the pixel, corresponding to the radiation absorption of the respective voxel from the anatomical object).

COLOUR MODELS IN IMAGE TRANSMISSION AND PROCESSING

The transfer from black/white (B/W) to colour television in mid-20th century led to a different way of colour representation. The TV engineers at that time decided that if they would transfer separately a signal related to light intensity and signal related to the colour, a TV monitor could reproduce respectively either B/W image or colour image. This is how another colour model was introduced – YUV, where the light intensity Y (aka "luma" signal, used for B/W images) was separated from the two colour (aka "chrome") signals U and V. This separation, applied to analogue TV transmission, used information from all three RGB colour signals to form the Y signal and a specific formula to form U and V [18]. The matrix below shows a typical conversion from RGB to YUV – here Y varies from 0 to 1 (as RGB), while U and V vary from minus to plus values (U: -0.436 to +0.436 and V: -0.615 to +0.615) [20].

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The YUV colour model has been used for analog PAL and analog NTSC TV signals (in some cases the model have some variations in calculation of the three components). Importantly the YUV model allowed reduction of the bandwidth in the TV transmission of the two "chroma" signals – i.e. allowed transmission of reduced number of colours, without significant compromise of the colour perception of the eye [19].

The introduction of digital imaging opened new horizons to image processing, compression and transmission. The necessity of more colours in the imaging information required a new presentation of the colour space. This was solved by a new technical application of the artistic colour scheme. This breakthrough was realised by Dr Alvy Ray Smith (with degrees both in Arts and Computer Science), who introduced the HSV colour model in the mid-1970. This is how he describes the development of the HSV model [21]:

"One of my best friends showed me his paint program at this new place called Xerox PARC. I knew this was exactly what I was looking for, a combination of art and computers. I got hired on as an artist to show off Dick Shoup's SuperPaint program (software and hardware). Now I can tell you the story of HSV.

Early in my encounter with SuperPaint I went to Dick and asked him what the algorithm was for converting the natural video color space of RGB to HSB (as I called it at first and as it is still called in Photoshop), which we artists use. I told him that I couldn't figure out how to make pink, say, with RGB, but for artists it was easy: choose red paint, add white paint to it to lighten it to pink. Easy! I also used brown as another example. Again I couldn't figure out how to make brown from RGB. But for artists it's easy: choose an orange red hue and darken it with black. Easy! That's when Dick informed me that no such algorithm existed.

So I started working on the algorithm. It had to exist obviously. So by the next morning I had it. I got to it by standing the RGB cube up on its main diagonal (the one from black to white), looking at it down that axis. That's where the hexagon comes from. The rest of it I just worked out so as to have a hue circle (like in a paint store), although the "circle" was a hexagon in this case. I thought first about a double-ended cone then rejected it as just a little more complex than a single-ended cone. (In those days every extra step cost.) That's it. What was wonderful about it was that it was VERY simple, which mattered a lot. So I coded it and inserted it into Dick's program. In short it took only a day to do the whole thing.

Then I spent years explaining to people that, despite its axes names, it wasn't REALLY a perceptually based system. That's when I stopped calling it Hue Saturation Brightness. The B axis was clearly NOT brightness. It was the amount of black, or Value as artists call it. And I insisted that it be called HSV after that."

The actual RGB to HSV transformation is described in SIGGRAPH 78 Conference Proceedings, Aug 1978 [23]. It should be mentioned that this transformation did not require expensive software techniques (very important in these early days of image processing).

The transfer of the artistic understanding of colours into the technical world revolutionised digital imaging as, although the immediate visual result is not intuitive, it allows new ways for image compression where significant reduction of image file size is achieved with minimal disturbance of the visual perception. This new concept founded the digital paint systems and opened new ways for digital image processing.

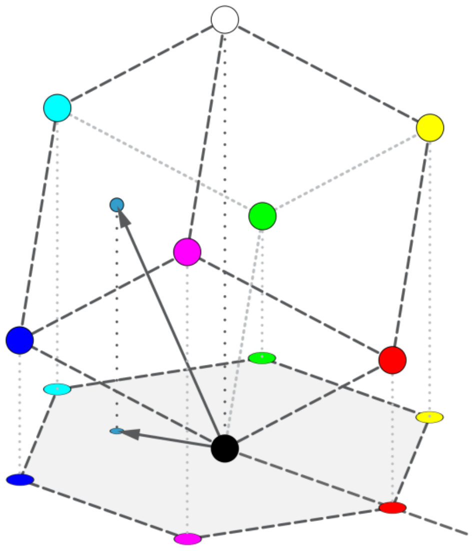


Figure 3. The figure is used to support Dr A R Smith description - projection of the RGB cube (with its CMY values, opposite their respective “additive” colours) into a hexagon (image cited from [22])

In 1982 the TV colour model also moved to digital with the YCbCr presentation of colour space, defined in the ITU-R BT.601-5 and ITU-R BT.709-5 standards of the International Telecommunication Union (ITU) [24]. YUV and YCbCr are similar, but YCbCr applies to digital systems, while YUV is for analogue TV systems.

COLOUR MODELS AND DIGITAL IMAGE COMPRESSION

When the RGB model is used, each colour could be presented in a coordinate system, where each of the “primary” RGB colours varies from zero to maximum value (e.g. from 0 to 1, or from 0% to 100%, or from 0 to 255 brightness levels, etc). The method is precise but requires lots of memory (full size of the image file) to describe the RGB coordinates of each colour from the light spectrum. Additionally, the RGB (or its subtractive CMY model) is not intuitive (i.e. not exactly related to the way humans perceive a colour). The image below (Figure 4) illustrates the difference between the two presentations of colour in RGB and HSV models, resulting from using Matlab for

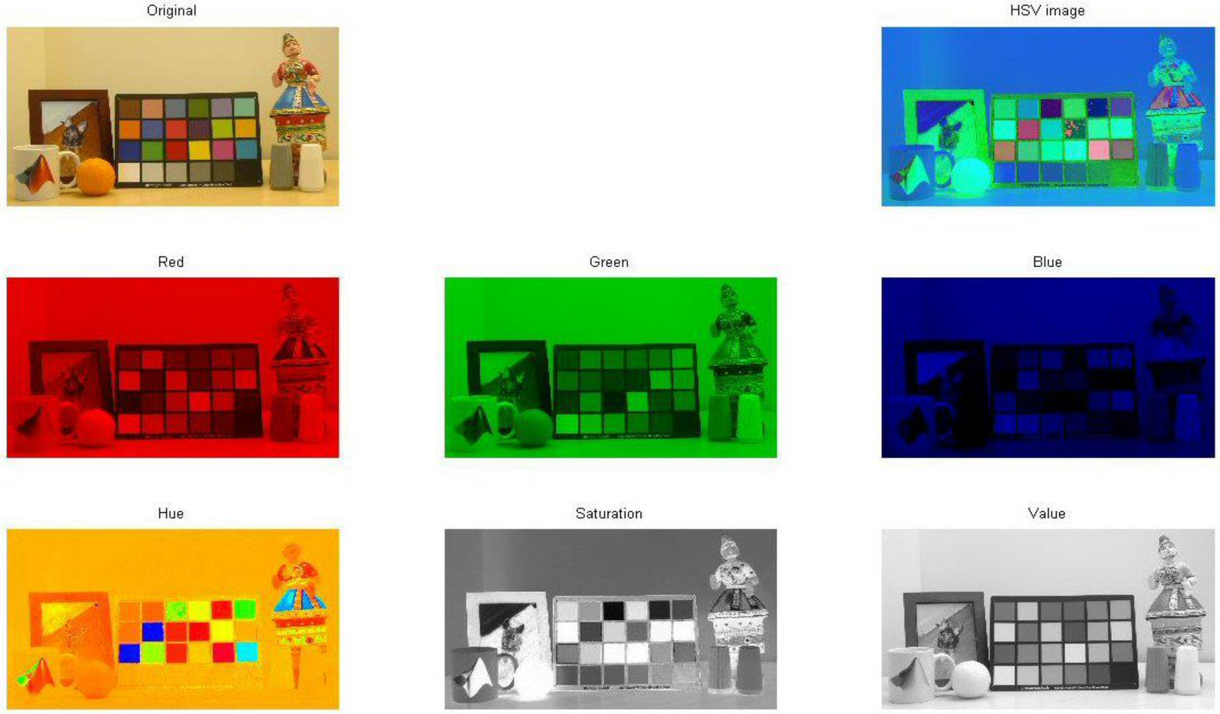


Figure 4. Illustration of the use of MatLab for decomposition of colour image – original image is decomposed to RGB and HSV (image courtesy to Dr A De Stefano).

decomposition of image colour [25].
The HSV model applies a measure of the light intensity (Volume – V); measure of the colour wavelength (Hue – H); measure of the amount of colour (Saturation – S). This model, although also not intuitive, is closer to the way of human perception and understanding of colours [25] – see Figure 5. Due to this it forms the background of the artistic system for colour production by mixing of paints. Even so, the HSV is usually not applied as an absolute colour model, but mainly as a method to encode the RGB information .

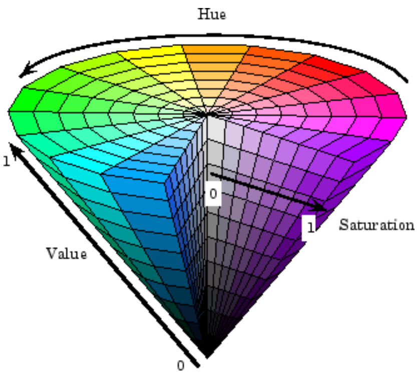


Figure 5. The cone of the HSV Colour Space (image cited from [26])

As explained above, the human eye is very sensitive to changes of image brightness (light intensity, Value), while it is less sensitive to change in chrominance. This means that, depending on the image, the visual data includes redundant information (psycho-visual redundancy) [27]. This redundancy allows to present the eye with full information related to Intensity (V), and with only limited information about the Hue and Saturation (usually only half of the "chroma" channels). This reduction of visual information does not change significantly the visual perception of the colours, but the image file size is significantly reduced. This way the colour model with separate Intensity/Value and "Chroma" channels forms the base of various image compression algorithms, which include certain loss of visual information.

The main compression steps of one of those widely used lossy algorithms – JPEG [28, 29] – includes (as per Figure 6) transforming the RGB image components into HSV (or similar colour model as YCbCr), thus forming three image files. These files are treated differently – the Value/Intensity file remains unchanged, while the two "chroma" files are reduced to half volume each (colour downsampling). The image components are then divided into small sections (e.g. 8x8 pixels) and each section is transformed in the frequency space through Direct Cosine Transformation (DCT). This transformation separates the high and low spatial frequencies in the image, thus allowing at the next stages (Quantisation and Encoding) some high frequencies to be discarded without significant loss of spatial resolution. This allows for application of user-defined 'image compression quality factor' (removing different percentage of the high spatial frequencies) – the higher the quality factor, the more high spatial frequencies will be preserved (but the larger will be the file). During the encoding the system selects the accuracy with which image components with different spatial frequencies are presented (e.g. low frequency components, which bring more visual information, are stored with higher accuracy). This final stage reflects the specificity of human vision to detect small variation in light intensity (brightness) over large areas (i.e. at lower spatial frequencies) [30].

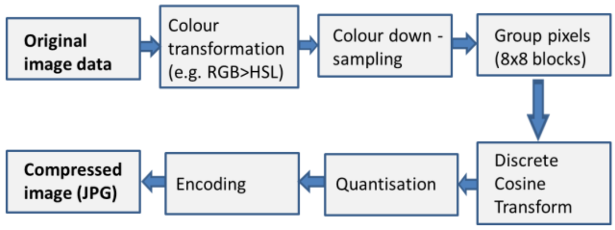


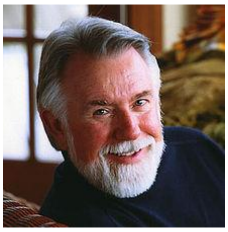
Figure 6. Simplified steps of the JPEG compression process

One very important moment in the compression process is to distinguish the relevant information, which should be preserved and this is a very difficult process, especially for medical images where clinically useful data must be carried with high resolution [31]. Once the compressed image is transmitted (or stored) with significant reduction of file size, its further re-visualisation requires all steps back through decoding and transformation into RGB in order to be presented to the final display monitor.

CONCLUSION

Different image compression algorithms apply different methods for reduction the image data file and not all algorithms are suitable for medical imaging. However all these methods are based on the fact that human vision has specificities which allow an image to deliver the necessary information with fewer resources. The process of transforming/encoding of the image colour space (resulting in acceptable reduction of information) requires good knowledge of the visual perception and imaging science, plus lateral thinking and creativity. Without a doubt the ideas behind the digital HSV colour model and algorithm are milestones of digital imaging. One field where these ideas are used is medical physics, but one could find them in various areas of life.

Information about human vision and perception of images/colours is essential for medical imaging, but this is a subject discussed in many different fields – from physiology to photography, and is rarely found in one place. Some useful References on the subject are given below. Elements from this Education Resource can be used in various introductory lectures to modules covering the physics of medical imaging.



Dr Alvy Ray Smith

Cofounded two successful startups: Pixar and Altamira (sold to Microsoft). Was present at the beginning of computer graphics at Lucasfilm and the New York Institute of Technology. Was the first Graphics Fellow at Microsoft. Received two technical Academy Awards for the alpha channel concept and for digital paint systems. Invented, directed, originated, or otherwise instrumental in the following developments: first full-color paint program, HSV (aka HSB) color model, alpha channel and image sprites, Genesis Demo in Star Trek II: The Wrath of Khan, first Academy-Award winning computer-generated short Tin Toy, first computer-generated film Toy Story, Academy-Award winning Disney animation production system CAPS, and the Visible Human Project of the National Library of Medicine. Was a star witness in a trial that successfully invalidated five patents that threatened Adobe Photoshop. Has PhD from Stanford University and honorary doctorate from New Mexico State University. Is a member of the National Academy of Engineering. Has published widely in theoretical computer science and computer graphics, and holds four patents. Retired in 2000 to devote time to the emerging artform of digital photography and to scholarly genealogy, to which he has contributed two award-winning books and half a dozen learned journal papers. He is Trustee Emeritus of the New England Historic Genealogical Society, Boston, and a Fellow of the American Society of Genealogists. He is now writing a book on the biography of the pixel. For more see <alvyray.com>.

REFERENCES

1. Helmholtz H, Treatise on Physiological Optics, English translation 1924, http://poseidon.sunyopt.edu/BackusLab/Helmholtz/
2. Munsell, A, A Pigment Color System and Notation, 1912, The American Journal of Psychology (University of Illinois Press) 23 (2): 236–244.
3. Wade N, Swanston N, Visual Perception: An Introduction, 2001, Psychology Press, Hove, East Sussex and Taylor and Francis, Philadelphia
4. Svaetichin G, Spectral response curves from single cones, 1956, Actaphysiol. Scand. 39, Suppl. 134, 17-46.
5. Wikipaedia, RGB Color model, (accessed 25 Feb 2013), http://en.wikipedia.org/wiki/RGB
6. Bayer B, Colour Imaging Array, U.S. Patent No. 3971065 (filed 5 Mar 1975)
7. Wald G, Brown P, Visual Pigments in Single Rods and Cones of the Human Retina, 1964, Science 144: 45–52
8. MacNichol E, Three Pigment Color Vision, 1964, Scientific American, vol. 211, No.6, 48-56
9. Wyszecki G, Stiles W, Color Science: Concepts and Methods, Quantitative Data and Formulae, 2000, Wiley-Interscience, New York
10. Wikipaedia, Rod cell, (accessed 25 Feb 2013), http://en.wikipedia.org/wiki/Rod_cells
11. Bowmaker J, Dartnall H, Visual pigments of rods and cones in a human retina", 1980, J. Physiol. 298: 501–511
12. Rods and Cones, (accessed 28 Feb 2013), http://hyperphysics.phy-astr.gsu.edu/%E2%80%8Chbase/vision/rodcone.html
13. Hecht E, Optics, 2nd Ed, 1987, Addison Wesley, Reading, MA
14. Tabakov S, Lecture Notes PACS – part 2 Viewing, 2009, King's College London
15. Barten P, Contrast Sensitivity of the Human Eye and its Effects on Image Quality", 1999, SPIE, Washington
16. Elert G (Editor), The Physics Factbook, Number of Colors Distinguishable by the Human Eye, (accessed 28 Feb 2013), http://hypertextbook.com/facts/2006/JenniferLeong.shtml
17. Kimpe T, Tuytschaever T, Increasing the Number of Gray Shades in Medical Display Systems—How Much is Enough?, J Digital Imaging, 2007 December; 20(4): 422–432
18. Tabakov S, Milano F, Strand S-E, Lewis C, Sprawls P (Editors), Encyclopaedia of Medical Physics, 2013, EMITEL Consortium and CRC Press, Boca Raton, FL
19. Wikipaedia, YUV, (accessed 3 March 2013), http://en.wikipedia.org/wiki/YUV
20. Equasys, Colour Conversion, (accessed 3 March 2013), http://www.equasys.de/colorconversion.html
21. Smith A R, 11 Nov 2011, Personal communication
22. Wikipaedia, HSL and HSV, (accessed 7 March 2013), http://en.wikipedia.org/wiki/HSL_and_HSV
23. Smith A R, Color Gamut Transform Pairs, SIGGRAPH 78 Conference Proceedings, Aug 1978, 12-19, available from: http://alvyray.com/Papers/CG/color78.pdf
24. ITU Standards, (accessed 3 March 2013), http://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.601-5-199510-S!!PDF-E.pdf
25. Tzu Yen's Matlab Codes, (accessed 9 March 2013), http://www.csse.uwa.edu.au/~wongt/matlab.html
26. MathWorks, Converting Color Data Between Color Spaces, (accessed 9 March 2013), http://www.mathworks.co.uk/help/images/converting-color-data-between-color-spaces.html
27. Pennebaker, W. B. & Mitchell, J. L. 1993. JPEG still image data compression standard, Norwell, Massachusetts, Kluwer Academic Publishers pp. 1-4
28. Milano J, Compressed Image File Formats: JPEG, PNG, GIF, XBM, BMP, 1999, Addison Wesley, Boston
29. Pennebaker W, Mitchell J, JPEG: Still Image Data Compression Standard, 1992, Kluwer Academic Publishers, Massachusetts
30. Kajiwara K, JPEG compression for PACS, Computer Methods and Programs in Biomedicine, 1992, 37, 343-351
31. Kim Y, Horii S, Handbook of Medical Imaging, Volume 3 - Display and PACS, 2000, SPIE Press, Bellingham
32. Wikipaedia, Bayer filter, (accessed 10 March 2013), http://en.wikipedia.org/wiki/Bayer_filter

Corresponding author:

Author: Slavik Tabakov

Institute: Dept. Medical Eng. and Physics, King's College London

Street: Denmark Hill

City: London

Country: UK

Email: slavik.tabakov@emerald2.co.uk