Color Image Quality for Multi-Function Peripherals

Jon Y. Hardeberg
Gjøvik University College
Gjøvik, Norway

Abstract

Color image quality is becoming an increasingly important factor in the consumer imaging industry. Users of imaging devices such as Multi-Function Peripherals (MFP) have increasing expectations to the quality of the reproduced images. In this paper we address the subject of color image quality from a practical point of view, and from the point of view of a provider of imaging technology for consumer MFPs. We show how the notion of color image quality is ultimately tied to the preferences of the end users. Because of this, practical quality evaluation experiments involving a panel of human observers is a very useful tool to quantify color image quality. As an illustration, we then describe a color image quality evaluation experiment, which was carried out in order to benchmark the copy function of two MFP devices.

Introduction

In the last two decades we have seen the field of digital color imaging emerging from specialized scientific applications, into being a part of the daily lives of most people in industrialized countries. Broadcast television, computers, newspapers and magazines, are just a few examples of technologies and media relying heavily on digital color imaging.

The increased use of color has brought with it new challenges and problems for the imaging technology providers. People's quality requirements have increased considerably. Just a few years ago, a computer system capable of producing 256 different colors on the display was more than enough for most users, while today, practically all computers that are sold have *true color* capabilities, being able to produce 16.7 million colors. Within the market segment known as consumer imaging, the transition from black and white to color has been very rapid in the last couple of years. Now, it would be difficult to purchase a black and white desktop scanner or inkjet printer.

Color image quality is thus becoming an increasingly important factor in the consumer imaging industry. Users of imaging devices such as MFPs have increasing expectations to the quality of the reproduced images. A company providing imaging technology is then faced with several questions related to color image quality: How do we optimize the color image quality delivered by our products? How do we judge color image quality? What does a good image look like? What are the best values in the color look-up tables (LUT)? Are our products better than our competitor's? Is the extra cost of using a particular technology justified in terms of increased quality? These questions are not easy ones, and we do not aspire to provide a straightforward recipe for answering them all here. It is, however, our hope that this paper can be a help in pointing out some of the important factors, and emphasize the need for conscious decisions about color image quality.

One of the conclusions that was drawn during the MFP Focus Group Sessions in Chicago in June, 2000[1] was that one of the most important deciding factors when customers compare different MFPs before buying, is the color image quality of the demonstration prints. Being able to produce images of high quality is therefore very important. And to make this happen, it is important to be able to evaluate and quantify color image quality.

After a brief description, in the next section, of what an MFP is and how it is created, we give an overview of the concept of color image quality. In the following section we describe a color image quality evaluation experiment, and finally we discuss the results and draw some conclusions.

MFP Technology Background

For the sake of the example, we consider here the following typical business scenario. A company desires to produce and sell an MFP. After having done the necessary market research, established the desired specifications, etc., it then assembles the MFP from parts purchased from different Original Equipment Manufacturers (OEM); the scanner system from one company, the print engine from another, the controller chip, the firmware, the software, the color look-up tables (LUT), etc., from yet other companies. Of course, one company might deliver several of these components, or the first company might do everything in-house, but typically, several companies are involved in this process.

The controller chip has several functions in the device, such as communicating with the I/O-devices, and performing the different image processing algorithms required for color copy. Particularly relevant to the subject of this paper is the function color conversion, which is typically based on 1D and 3D LUTs (Figure 1). Different LUTs are needed for different operations such as color copy, print from host, color fax, etc. Optimization of the LUTs using techniques such as colorimetric device characterization[2–4] is an important step towards achieving high color image quality in the final product.

According to the principles of color management, the content of the LUTs depends on the actual I/O-devices that is being used for a given MFP, but as we will see, also on a great deal of more or less subjective judgements concerning color image quality. The focus of this paper will be on the aspects of color image quality that are influenced by the color conversion, in particular on the actual content of the LUTs.

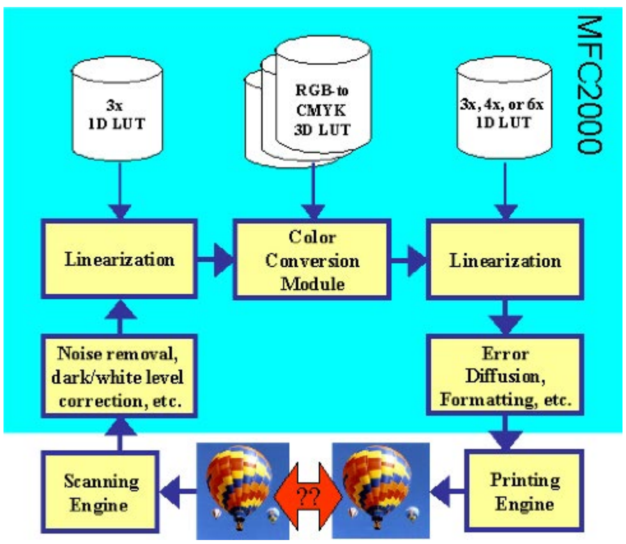

Figure 1. Simplified flowchart of the color copy operation in ViewAhead's MFC2000 MFP controller chip. The color LUTs are optimized to achieve the best color image quality in the copied image.

Color Image Quality

As mentioned in the introduction, color image quality is of very high importance in a digital imaging device such as an MFP. For a manufacturer of imaging technology it is therefore important to be able to quantify color image quality. Potential uses of quantifiable data on color image quality include:

- Tradeoff analysis of speed and implementation cost versus color image quality in image processing algorithm development.
- Benchmarking of imaging systems and algorithms to other vendor's products.
- Documentation of color image quality improvements resulting from efforts spent on optimization of image processing algorithms.

In recent years, the concept of image quality has received quite much attention within the imaging science and technology community. In particular, the subject has been extensively discussed at the PICS conferences.[5–11] But still, image quality often receives a rather stepmotherly treatment in the industry — probably because of its somewhat awkward position between subjectivity and objectivity.[7] The concept of quality, typically defined in dictionaries as 'degree of excellence' is inherently a subjective entity. An engineer and scientist, however, generally prefers to deal with objective quantities, backed by scientific evidence.

On the web site of a major French consumer electronics retailer, a formula for the image quality of a printer was given approximately as follows: *Image quality = Resolution x Color Depth*. This is an example of another common misconception regarding image quality — its oversimplification. There are indeed many factors that contribute to the quality of an image, such as spatial resolution, color depth, the "nesses" (sharpness, naturalness, colorfulness, etc.), and visual artifacts (banding, streaking, grain, blocking, mottle, moiré, etc.). There exist an ongoing effort to standardize the definitions of these and other image quality factors, as well as their assessment methodology, see for example a recent paper by Grice and Allebach.[12] Further discussion of this topic is, however, beyond the scope of this paper.

We will therefore proceed to a discussion of color image quality factors with direct relevance to the creation of color LUTs, such as the uniformity of color gradations (sweeps), and color reproduction accuracy. We then describe the measurement of color differences such as ΔE^*_{ab} as a tool to quantify color quality, in particular we discuss the limitations of this approach. Finally, based on the realization that the ultimate definition of color image quality is what the end users prefer, we present a very important way to quantify color image quality in practice—to perform quality evaluation experiments with several human observers.

Color Image Quality Factors

The content of the color LUTs that define the different color conversions is of particular importance to the color image quality of an MFP. Here, we pay special attention to the color copy function, which typically require a color conversion from the RGB color space proper to the scanner, to the CMYK color space of the printer. To a certain extent, the content of the relevant LUTs is determined by the hardware, but it is important to realize that there is a large room for variation, and that many parameters, design criteria, and trade-offs should be considered.

For example, it is impossible for the device manufacturer to control which media the end user will use with the device, in particular the paper type used for printing. It is common industry practice to categorize paper into a few classes, such as plain paper, coated paper, and glossy paper, to provide specific LUTs for each of these classes, and to expect the end user to specify the right paper type in a user interface to the device. But unfortunately there are also significant variations of the reproduced color between different papers within the same class.[13]

The two quality factors of a color LUT that are probably the most important can be labelled as "correctness" and "smoothness." Correctness obviously means that the LUT should contain the right colors, but to define exactly what those right colors should be is not that easy. Typical goals for an MFP are that the colors of a reproduction should be exactly the same as in the original, and that what you see on the monitor is what you get printed out on paper. However, because of limitations in the printer color gamut it is mostly impossible to achieve this correctness for all colors—trade-offs must be done. It is for example of particular importance that colors that are neutral gray in the original remain neutral in the reproduction. Skin tones are also very important to reproduce correctly.

Smoothness means that there should be no abrupt steps in the LUTs — the values should vary smoothly in the entire LUT. Violating this typically leads to visual artifacts such as banding and 'false contours,' especially visible in so-called color ramps (also known as gradations or sweeps). These typically occur in backgrounds of business graphics, but also in natural scenes, see Figure 2. Olson[14] has discussed this issue. Especially when the gamut differences are large, it is necessary to perform a trade-off between correctness and smoothness. A good gamut mapping algorithm need to consider this tradeoff.[15]

Figure 2. A natural color image processed through two different ICC profiles. Banding artifacts due to a "bad" profile, are apparent in the background of the left picture. (Illustration courtesy of Olson[14].)

Color Difference Measurements

One means of evaluating color quality is through the measurement of color differences between the actual reproduction and a preferred color reproduction. The preferred color reproduction typically relates to an original document when evaluating color copy, the colors as they appear on the monitor for a typical "WYSIWYG" evaluation, or specific colors of which the color reproduction is particularly important for a given reproduction.

The color differences are typically measured in terms of ΔE^*_{ab} — the Euclidean distance between two colors in the CIELAB color space.[16,17] Since natural images rarely contain sufficiently uniform areas to allow consistent color measurements, a color target with several uniformly colored patches should be used. Simple statistical measures such as maximum and average color differences are then typically used as an indication of the color quality. Note that newer formulae for color difference are slowly replacing the ΔE^*_{ab}, the most recent one being CIEDE2000.[18]

A very helpful resource for such color quality evaluation is the Microsoft Windows Color Quality Test Kit.[19] This freely available kit contains descriptive documents, test targets and images, as well as tools for the calculation of color differences, for several different color imaging devices. Typically the goal is for the devices to communicate images using the sRGB color space.[20,21] If certain criteria, in particular in terms of average color difference, are not met, the device does not receive Microsoft's certification — the 'designed for Windows' logo.

As a final note on color quality evaluation through color difference measurements, we would like to remark that the numbers should be used with care. Although being a very good indicator for color quality, minimizing the average and maximum color differences does not guarantee optimal results in terms of perceived color image quality. For example, colors that are not in the evaluation target might be important. Another factor that limits this approach is gamut mapping.[15] Because of the differences in color gamut between different devices and technologies, a colorimetrically exact reproduction is rarely optimal.

Color Quality in Practice — Customer Preference

In slight contrast with the traditional definition of quality as 'degree of excellence', the standardized definition of quality in ISO 9000 refers to all those features of a product (or service) which are required by the customer.[22] Color image quality is what the customer wants! The typical challenge an imaging company is faced with is how to optimize certain technology variables so as to maximize the color image quality as judged by the customer — this is represented by the thick arrow in Figure 3.

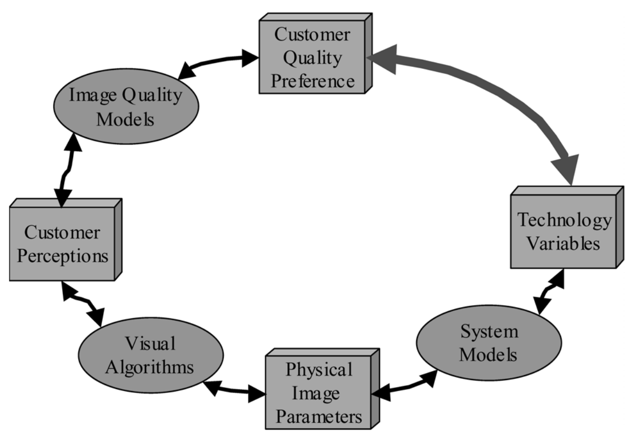

Figure 3. The Image Quality Circle. The desired mapping between the technology variables and the customer's perception of quality is sought to be found through a three-level approach — system models, visual algorithms, and image quality models. Adapted from Engeldrum.[23]

In practice, the way this often works in an OEM customer relationship, is that you try something and present it to the customer. If the customer accepts, you're OK, but if not, you'll have to try to understand what the customer means by for example "the reds are no good" and start the loop again. This is a very time-consuming and potentially frustrating procedure.

The ideal situation would be if there existed a formula mapping the technology variables you are trying to optimize, to a quantification of color image quality. Much research effort is being spent on this task, see for example the book of Engeldrum,[23] but we feel that so far, the proposed solutions are not ready for industrial 'prime time.'

Ultimately, color quality is a question of preference. Some people tend to prefer "warmer" colors than other, for example. American movies are generally much more colorful than British ones. The same printer is sold with different color profiles in Japan and USA. Knowing this, the challenge is to enable the optimization of the various LUTs and parameters that influence the color reproduction without relying too heavily on the personal preferences of the color engineer who is doing the work of creating the LUTs and optimizing the image processing parameters.

A very important tool to this end is to perform psychophysical experiments with a set of human observers. The observers are asked questions relating to the quality of images reproduced using different technologies, algorithms, and/or parameters, and the answers are analyzed statistically, as described in the next section.

Experimental Results

In this section we present the results of a color image quality evaluation experiment which was carried out in order to benchmark two MFP devices from different manufacturers. We were particularly interested in the quality of the LUTs used for color conversion in the copy function.

A color image quality evaluation experiment was carried out. It involved 7 observers. Each observer was asked to evaluate 21 printed images. The images were 3 originals (Figure 4), and each of these were copied on 2 different MFP devices at 3 different quality levels (draft mode on plain paper, normal mode on plain paper, and normal mode on glossy paper). The images were labeled using codes of the form **I***N*-*K*, where $N = 1, 2, 3$ refers to the motif/original image, while $K = 0$ refers to the different copy methods described above.

For each of three questions (Table 1) representing three quality factors (The pleasantness of the colors, Overall Image Quality, and color match to the original), they were asked to pick one answer out of seven alternatives: Excellent, Very Good, Good, Fair, Neither/Nor, Poor, Very Poor.

The images were viewed under Cool White Fluorescent light, using a GretagMacbeth SpectraLight 3 light box.

Figure 4. The three images used in the experiment. Top: ***I1****, a test target provided by an MFP manufacturer. Left:* ***I2****, a composite image created using SCID24 images, printed on a continuous-tone Fujix Pictography printer. Right:* ***I3****, a mixed-content page taken from a magazine.*

Table 1. **Questions asked to the observers.**

Q1: How would you characterize the *pleasantness* of the *colors* of this print?
Q2: How would you characterize the *overall image quality* of this print?
Q3: Identifying prints **I1-0**, **I2-0** and **I3-0** as originals, and the other prints with the same motif as copies, how good would you say the *color match* is between original and copies?

All samples were compared blindly. No apriori knowledge of print technology, quality modes, or paper-types were given to the observers. For Q1 and Q2 they were explicitly asked not to consider whether a print was an original or a copy, while for Q3 the original prints were identified. The images were presented in a consistent pseudorandom sequence, and the observers were allowed to make corrections/revisions to their answers during the procedure.

To analyze the observer ratings, each of the possible answers was assigned an Image Quality Score value, ranging from 0 for Excellent, to 6 for Very Poor. We report in Table 2 and Figure 5 the average IQ scores for each device/mode, along with an attempt to interpret these results on a scale from Excellent to Very Poor.

Table 2. Color Image Quality Evaluation Experimental Results

Q1: Color Pleasantness	**Score**	**Rank**	**Category**
Device#1 Plain Draft	3.67	5	Neutral
Device#2 Plain Draft	3.67	5	Neutral
Device#1 Plain Normal	3.11	4	Fair
Device#2 Plain Normal	2.50	3	Good
Device#1 Glossy Normal	1.89	1	Good
Device#2 Glossy Normal	2.00	2	Good
Q2: Overall Image Quality	**Score**	**Rank**	**Category**
Device#1 Plain Draft	4.33	6	Neutral
Device#2 Plain Draft	3.94	5	Neutral
Device#1 Plain Normal	3.28	4	Fair
Device#2 Plain Normal	2.67	3	Fair
Device#1 Glossy Normal	1.94	1	Good
Device#2 Glossy Normal	2.50	2	Good
Q3: Color Match	**Score**	**Rank**	**Category**
Device#1 Plain Draft	2.72	5	Fair
Device#2 Plain Draft	3.17	5	Fair
Device#1 Plain Normal	3.00	4	Fair
Device#2 Plain Normal	2.56	3	Fair
Device#1 Glossy Normal	1.94	1	Good
Device#2 Glossy Normal	2.39	2	Good

Comparing the two devices, we see that for glossy paper and normal quality, Device#1 gives better quality, while for plain paper and normal quality, Device#2 outperforms Device#1. In draft mode, Device#2 has better overall image quality, Device#1 has better color match, while the color pleasantness is judged to be approximately equal on the two devices.

If we analyze the results per quality mode, we see that as expected, for normal mode, glossy paper gives better quality than plain paper, and for plain paper, normal mode is better than draft mode. One exception to this is the color match of Device#1, which is better for draft mode than for normal mode plain paper.

From the results, we can also conclude that there is more correlation between color pleasantness and overall image quality, than between color match and overall image quality, that is, to achieve high image quality, it is more important to aim for pleasing colors than colors that match the original.

Discussion and Conclusion

Color image quality is of very high importance in a digital imaging device such as an MFP. For a manufacturer of imaging technology it is therefore important to be able to quantify color image quality. However, to do so is not a trivial task, since ultimately, quality is defined as what the customer wants. Unfortunately, as of today there are no analytical techniques that can quantify color image quality in this context. It is therefore necessary to rely on experiments involving real observers, and such experiments has been shown to be very useful in the product development process.

However, it is clear that such experiments are relatively time consuming. Definitively, Yendrikhovskij[7] hits the nail on the head when he states that "most studies on image quality employ subjective assessment with only one goal — to avoid it in the future." Therefore results from ongoing research toward analytical models for color image quality is eagerly anticipated. An example of such research is the development of metrics for color differences between complex images.[25,26] However, a device or algorithm that takes any image as input, and provides a number that pefectly quantifies its color image quality as output, is still probably many, many years away.

As a final note, we mention that in *practice*, in the MFP imaging industry, color image quality is not primarily judged by the end user. Typically, the color LUTs are produced by an OEM, and it is representatives from the company that produces the MFP who decide when the quality of the color LUTs is acceptable. Unfortunately, their preferences do not necessarily match those of the end users...

Acknowledgments

The described image quality evaluation experiment was carried out while the author was employed with Conexant Systems/ViewAhead Technology. The participation of my former colleagues as observers, and of the management by facilitating this paper, is highly appreciated.

References

1. Flatbed MFP and Focus Groups. Report Presented by Priorities Consumer Insights in cooperation with CCSI Market Intelligence Department, Canon Computer Systems, Inc., July 2000.
2. Jon Y. Hardeberg. Desktop scanning to sRGB. In *Color Imaging: Device Independent Color, Color Hardcopy and Graphic Arts V*, volume 3963 of *SPIE Proceedings*, pages 47–57, 2000.
3. Jon Y. Hardeberg. *Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches*. dissertation.com, Parkland, Florida, USA, 2001. ISBN 1-58112-135-0. Available at http://www.dissertation.com/library/1121350a.htm. (Revised second edition of Ph.D dissertation, Ecole Nationale Supérieure des Telecommunications, Paris, France, 1999).
4. Ying Noyes, Jon Y. Hardeberg, and Anatoly Moskalev. Linearization curve generation for CcMmYK printing. In *Proceedings of IS&T and SID's 8th Color Imaging Conference: Color Science, Systems and Applications*, pages 247–251, Scottsdale, Arizona, 2000.

5. Michael Stokes. The impact of color management terminology on image quality. In *Proc. IS&T's 1998 PICS Conference*, pages 174–178, Portland, Oregon, 1998.
6. D. René Rasmussen, Peter A. Crean, Fumio Nakaya, Masaaki Sato, and Edul N. Dalal. Image quality metrics: Applications and requirements. In *Proc. IS&T's 1998 PICS Conference*, pages 216–220, Portland, Oregon, 1998.
7. Sergej N. Yendrikhovskij. Image quality: Between science and fiction. In *Proc. IS&T's 1999 PICS Conference*, pages 173–178, Savannah, Georgia, 1999.
8. Karin Töpfer and Robert Cookingham. The quantitative aspects of color quality rendering for memory colors. In *Proc. IS&T's 2000 PICS Conference*, Portland, Oregon, 2000.
9. Paul J. Kane, Theodore F. Bouk, Peter D. Burns, and Andrew D. Thompson. Quantification of banding, streaking and grain in flat field images. In *Proc. IS&T's 2000 PICS Conference*, Portland, Oregon, 2000.
10. J. Jung, R. Geelen, and D. Vandenbroucke. The virtual image chain: A powerful tool for the evaluation of the perceived image quality as a function of imaging system parameters. In *Proc. IS&T's 2001 PICS Conference*, Montréal, Québéc, Canada, 2001.
11. Peter G. Engeldrum. Psychometric scaling: Avoiding the pitfalls and hazards. In *Proc. IS&T's 2001 PICS Conference*, pages 101–107, Montréal, Québéc, Canada, 2001.
12. Jim Grice and Jan P. Allebach. The print quality toolkit: An integrated print quality assessment tool. *Journal of Imaging Science and Technology*, 43(2):187—-199, 1999.
13. David L. Lee and Gabriel G. Marcu. Performance of media-specific and device-specific output profiles. In *Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts V*, volume 3963 of *SPIE Proceedings*, pages 119–130, 2000.
14. Thor Olson. Smooth ramps: Walking the straight and narrow path through color space. In *Proceedings of IS&T and SID's 7th Color Imaging Conference: Color Science, Systems and Applications*, pages 57–64, Scottsdale, Arizona, 1999.
15. Ján Morovic. *To Develop a Universal Gamut Mapping Algorithm*. PhD thesis, Colour & Imaging Institute, University of Derby, October 1998.
16. *Colorimetry*, volume 15.2 of *CIE Publications*. Central Bureau of the CIE, Vienna, Austria, 2 edition, 1986.
17. Günter Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley & Sons, New York, 2 edition, 1982.
18. Hue and lightness correction to industrial colour difference evaluation. CIE Publication No. TC1.47 — Tentative CIEDE2000 Color Difference Equation, 2000.
19. Windows Color Quality Specifications for Printer OEMs. Microsoft Corporation, 2001. Part of Microsoft Hardware Quality Labs (WHQL)'s Windows Color Quality Test Kit. http://www.microsoft.com/hwdev/tech/color/ColorTest.asp.
20. Matthew Anderson, Ricardo Motta, Srinivasan Chandrasekar, and Michael Stokes. Proposal for a standard default color space for the internet — sRGB. In *Proceedings of IS&T and SID's 4th Color Imaging Conference: Color Science, Systems and Applications*, pages 238–246, Scottsdale, Arizona, November 1996. Updated version 1.10 can be found at http://www.color.org/sRGB.html.
21. Colour measurement and management in multimedia systems and equipment, Part 2-1: Colour management – Default RGB colour space - sRGB. *IEC Publication*, 61966-2.1, 1999.
22. ISO 9000 and ISO 14000 in plain language. Online: http://www.iso.org/iso/en/iso9000-14000/tour/plain.html, 2001.
23. Peter G. Engeldrum. *Psychometric Scaling: A toolkit for imaging systems development*. Imcotek Press, Winchester, Massachusetts, 2000.
24. Graphic technology — Prepress digital data exchange — CMYK standard colour image data (CMYK/SCID). *International Standard ISO*, 12640, 1997.
25. Methods to derive colour differences for images. Draft Version 0.4 of CIE Technical Report, 2000. Online: http://www.colour.org/tc8-02/cietc802v04.pdf.
26. F. H. Imai, N. Tsumura, and Y. Miyake. Perceptual color difference metric for complex images based on Mahalanobis distance. *Journal of Electronic Imaging*, 10(2):385—393, 2001.

Biography

Jon Y. Hardeberg received his Ph.D. from the Ecole Nationale Supérieure des Télécommunications in Paris, France in 1999. His Ph.D. research concerned color image acquisition and reproduction, using both colorimetric and multispectral approaches. He then worked for 2.5 years as a color scientist with ViewAhead Technology (a.k.a. DeviceGuys, Conexant) in Bellevue,Washington, USA. He is currently Associate Professor with Gjøvik University College in Norway, where he is teaching and researching in the field of color imaging science. He is also part-time researcher with SINTEF Electronics and Cybernetics in Trondheim, Norway. Email: jon@hardeberg.com.