Høgskolen i Gjøviks rapportserie, 2011 nr. 6

Proceedings from Gjøvik Color Imaging Symposium 2011

Jon Yngve Hardeberg (ed.)

Gjøvik 2011

ISSN: 1890-520X ISBN: 978-82-91313-78-8

Table of Contents

Preface

For the sixth time, in what has become a biannual tradition, Gjøvik University College and The Norwegian Color Research Laboratory have organized an international symposium within the field of color imaging. Gjøvik Color Imaging Symposium 2011 took place in Gjøvik, Norway, on September 7-9, 2011. The symposium contained a rich program of tutorials, keynotes, invited, and contributed talks on a wide variety of color imaging research topics such as image quality, multispectral imaging, color management, color in printing, color and 3D, and color image processing, given by a balanced mixture of well known international experts and students in the field.

This year was a bit special, since we also took the opportunity to celebrate the Colorlab's ten year anniversary. A joint session on color imaging in collaboration with the 2nd International Conference on Pervasive Computing, Signal Processing and Applications (PCSPA 2011) was also organized. More information concerning the symposium can be found at http://www.colorlab.no/events/gcis11.

In these proceedings you will find abstracts and short papers corresponding to submitted and invited papers from the symposium.

Gjøvik, December 2011
Prof. Jon Y. Hardeberg, Symposium general chair

Visualization of Spectral Images

A Comparative Study

Steven Le Moan[1,2], Alamin Mansouri[1], Jon Yngve Hardeberg[2], Yvon Voisin[1]
1: Le2i, Université de Bourgogne, Auxerre, France
2: The Norwegian Color Research Laboratory, Gjøvik University College, Gjøvik, Norway

Abstract

The dimensionality reduction of spectral images for visualization has been a quite active area of research recently. Given the variety of existing approaches, it can be very challenging to understand the actual advantages of one over another, especially in the absence of a very specific application. Moreover, there is no consensus on how to evaluate the general efficiency of such a method. In this paper, we propose a comparison framework not only to compare such techniques, but also to measure their intrinsic properties in terms of naturalness and informative content.[1]

[1]The full paper will be published in Proceedings of "Second International Conference on Pervasive Computing, Signal Processing and Applications", September 2011, Gjøvik, Norway

Seam Carving for Multi-projector Displays

Marius Pedersen, Arne Magnus Bakke
Gjøvik University College
Gjøvik, Norway
{marius.pedersen,arne.magnus.bakke}@hig.no

Abstract

Improving the spatial and colorimetric models used in tiled display systems is a current research topic. Seam carving was recently proposed as an algorithm for intelligent content-aware resizing of images. We use an adapted version of seam carving for creating a less visible transition between two parts of an image, the two parts having been mapped to different color gamuts. The method is evaluated in a pair comparison experiment with 10 observers and 25 images, and is shown to perform significantly better than a transition along a straight line. Consequently, we suggest that seam carving can be used for improving the reproductions in multi-projector displays.[1]

[1]The full paper will be published in Proceedings of "Second International Conference on Pervasive Computing, Signal Processing and Applications", September 2011, Gjøvik, Norway

A Total Variation Based Color Image Quality Metric with Perceptual Contrast Filtering

Marius Pedersen[1,2], Gabriele Simone[1], Mingming Gong[1,3], Ivar Farup[1]
1: Gjøvik University College, Gjøvik, Norway
2: Océ Print Logic Technologies, Creteil, France
3: Huazhong University of Science and Technology, Wuhan, China

Abstract

In the last two decades, the evaluation between an original image and its reproductions has been widely considered by many researchers. Recent studies have shown that contrast is one of the most important image features falling under the umbrella of image quality factors. Total variation has shown to be a useful tool in different areas of computer vision. In this paper we introduce a novel image quality metric, named Total Variation of Difference (TVD), combining the total variation method with a local band-limited contrast filtering. Extensive tests and analysis of different pooling methods are carried out on two different databases. Results show a particular high correlation on the second database using Minkowski pooling.[1]

[1]The full paper will be published in Proceedings of "Second International Conference on Pervasive Computing, Signal Processing and Applications", September 2011, Gjøvik, Norway

Quality Improvement for Omnidirectional Spectral Images Captured with a Fisheye Lens

Takahiko Horiuchi, Shoji Tominaga, Shun Abe
Graduate School of Advanced Integration Science
Chiba University
Chiba, Japan
{horiuchi,shoji}@faculty.chiba-u.jp

Abstract

This paper describes a method for improving quality of omnidirectional spectral images in a natural scene. In our previous study, a multiband omnidirectional imaging system with a fisheye lens was proposed for capturing high resolution images. However, the fisheye lens enhances chromatic aberrations on the sensor plane. Moreover, since the fisheye system requires taking pictures of a scene in a hemisphere, we have to combine three sets of images observed at rotation angle intervals of 120 degrees. Therefore, the combined omnidirectional image produces certain geometric distortions at the edge of image planes. We propose compensation algorithms for eliminating the chromatic aberrations and the geometric distortions for omnidirectional spectral images. The feasibility of the proposed algorithm is tested using natural outdoor scenes.[1]

[1]The full paper will be published in Proceedings of "Second International Conference on Pervasive Computing, Signal Processing and Applications", September 2011, Gjøvik, Norway

Zen and the Art of Image Quality Assessment

An attempt to bridge the gap between subjective and objective image quality

Dag Waaler, Jon Yngve Hardeberg
Gjøvik University College
Gjøvik, Norway
{dag.waaler,jon.hardeberg}@hig.no

Abstract

The purpose of this paper is to examine whether the perspective in the novel "*Zen and the Art of Motorcycle Maintenance: an Inquiry into Values*", by Robert Maynard Pirsig, can shed new light on the meaning of the term *image quality*. The term is used in such different areas as medical imaging, broadcasting, printing industry, media production, marketing, graphics and art, to name but a few. The problem goes straight into a long standing philosophical debate about the differences between primary and secondary qualities, where primary qualities are thought of being inherent in an object, a thing or a person, and represent those aspects of objects that science can deal with, while secondary qualities depend on subjective interpretations and contexts. Pirsig seeks to reconcile these two seemingly irreconcilable visions of reality, and unites them into a greater unity: "Quality couldn't be independently related with either the subject or the object but could be found *only in the relationship of the two with each other*. It is the point at which subject and object meet... Quality is not a thing. It is an event."

1 Introduction

This paper is not about Zen. It is inspired, however, by the novel "*Zen and the Art of Motorcycle Maintenance: an Inquiry into Values*" by Robert Maynard Pirsig [14]. On the outside the novel is about a father and a son who, together with a couple of friends of his father on a motorcycle tour, which, eventually, proves to be a life odyssey about how we perceive the world depends on our intentions and commitments. Central in this are his reflections on the notion of quality, which Pirsig identifies as derived from two different and conflicting beliefs or understandings of reality, one of an immediate and artistic character (which he calls a "romantic world view"), and the other with emphasis on explanations and the scientific (as Pirsig calls a "classical world view"). And even if this is a fictitious novel, it is quite clear that the protagonist is the author's alter ego, and that the insight is based on his own experience and the conclusions are seriously meant. The purpose of this paper is to examine whether Pirsig's analysis can have something to say in the area of image quality. We emphasize, however, that we use the ideas quite freely.

2 Image quality

What is it that determines whether we consider an image to be "good" or "bad"? Taste? - Certainly! Technical details of the camera? - All right, that as well! Point of view, colors and lighting? Probably! Still further quality parameters can be e.g. how well an image can act as basis for decisions, how it is composed, if it possess the desired colors, do not show disturbing objects, succeed in conveying a specific mood, just to name a few.

As the above examples suggest, there are many things that may be important for the quality of an image. A typical way to clarify the situation is to introduce a distinction between

subjective and objective criteria. Composition, beauty, commitment, etc. are categorized as subjective criteria, on which we never can nor wish to agree upon. On the other side, the number of detector megapixels, lens features, image artifacts and noise are classified as objective criteria that can be defined unambiguously and thus objective and universal. Although this distinction at first thoughts seems to solve the problem, we will argue that this way of separating subjective and objective image criteria does not catch the essence of quality.

Googling the term "definition of image quality" returns a whole lot of suggestions of which but a few exceptions are of "objective" character. On further look, however, they are mostly related to imaging technology, not to the image as such, and also apparently without reference to any human observer. This preoccupation with "technical image quality" is relatively new, and probably stems from our extensive use of photos (and video, film) for measurement and classification (quality control, diagnosis, automatic sorting, etc.) in industry, medicine, surveillance, etc. and where precision is important. The scientific interest is also relatively new. One of the first people interested in this, Albert Rose, then at RCA (Radio Corporation of America). In the wake of the establishment of information theory and research on the radars in the 1940s Rose came up with a model of how human observers detect visual signals [6]. Particular interests for image quality in the photographic industry did not appear until the middle of the 1960s [7]. As a result, image quality in many contexts thus become synonymous with technical attributes such as sharpness, color fidelity, naturalness, noise, contrast, degradation, and even more sophisticated measures such as signal-to-noise ratio (SNR), detective quantum efficiency (DQE), Wiener spectra, etc. [13].

The incompleteness of defining image quality in terms of technical parameters is that it is implicitly linked to very specific tasks, particularly measurements and decisions. It is nothing wrong with this type of imaging tasks; the point here is just that a definition of image quality without reference to the particular imaging purpose has no meaning. A system that is optimized for one type of task may be suboptimal for another, and quite irrelevant to a third.

We have no ambition of digging deep into the concept of quality as such, but for our arguments sake a few words are necessary. The word quality is derived from Plato's "poiotes" can be translated as "what-ness" or "off-what-kind-ness" and was translated by Cicero to the Latin word "Qualitas" [2]. During the age of enlightenment the Greek dualism was taken a step further by philosopher such as René Descartes and John Locke. Of particular relevance here is the distinction between what Locke [11] defined as primary and secondary qualities. Primary qualities are inherent in an object, a thing or a person, and represent those aspects of objects that science can deal with, while secondary qualities depend on the subjective interpretation and the context. And this is the important point in our context: it is this dualistic way of thinking that Pirsig with its "*Zen and the art ...*" and later with "*Lila: an Inquiry into Morals*" [15] argues against.

A Google search on quality offers a wide range of definitions; ranging from the "subjectively" *quality is a degree of excellence* (Webster Dictionary), to the "objectively" *quality is a totality of characteristics of a product or a service that helps to meet certain needs* (American Society for Quality).

Pirsig seeks to reconcile the two seemingly irreconcilable visions of reality, i.e. the subjective, immediate and artistic expression on one side (a "romantic" view), on the other hand, the objective and scientific (the "classical" view). The classic view of reality is all about reason and evidence, often in the form of quantitative data and scientific methods to solve problems and answer questions. This strategy, however, transforms the quality to quantity (indicators), so it should come as no great surprise that those with a "romantic" conception of the world will find this uncomfortable, even meaningless. They will argue that quality in art, innovation,

imagination, etc. is something "qualitatively" (sic!) different, and impossible to measure. Whereupon the "classic" will respond that the "romantic" concept of quality is too vague and thus completely useless. This is where Pirsig enters with his integration of the two visions into a greater unity: "And by God, it wasn't subjective or objective either. It was beyond both *those* categories" (pp. 212-213), and "Quality couldn't be independently related with either the subject or the object but could be found *only in the relationship of the two with each other*. It is the point at which subject and object meet... Quality is not a thing. It is an event." (p. 215) (Italics in original)

3 Images

If Pirsig's view is a fruitful approach to the concept of quality it should also be relevant for image quality. And again, if applicable to both the romantic and the classical view, should it not also include all kinds of "images"? As previously mentioned, this article is primarily concerned with images that can be presented in digital form; it should also be interesting to consider how the term "image" (image quality) can be defined in a broader context. All images are purveyors of (images of) the "something". They are carriers of visual information, and as such important part of interaction with the environment [9], and in order to convey information they need interpretations. There is no neutral, univocal "visible world" there to match things against, no unmediated "facts" about what or how we see. Quite contrary, a necessary element in all interactions is purpose; there are no images (vision) without a purpose, a task, which could be of many kinds: performance, decisions, contemplation, aesthetics, joy, fear, propaganda, etc.

This illustrates another shortcoming of the many typical "definitions" of image quality, that they do not explicitly specify the observer, i.e. the interpretive authority, but implicitly assume that this is a human being. In addition to humans, however, other types of observers are also possible; e.g. animals, computers and theoretical model observers. A further division may be required as we know, for example, that experts and non-experts, experienced and beginners, children and adults, etc. do not necessarily evaluate equally. And we also know intuitively that a family picture is not considered in the same way by the photographer (who can be proud), the ones photographed (who might think they did not show their best side), and the rest of us (who may think the picture is quite boring).

4 Image quality research

A large part of the research in the area of image quality is related to applications where one wishes to establish objective measures of the relationship between how images are evaluated by observers (subjective) and the images of physical (technical) parameters. Examples include medical imaging [8, 16], broadcasting [4], and graphics / printing industry [12], to name but a few. The motivation for research in all these areas is the wish to automatically quantify given tasks (e.g. medical diagnostics, image transmission, color reproduction). Succeeding in simulating an observer panel by physical parameters that can be derived directly from the image itself will save both time and money. This feature is particularly desirable when the task is to optimize between several constraints. An example of this is X-ray diagnostics; because X-rays are harmful it is desirable to find the lowest dose that still provides adequate diagnostic decision accuracy. Another example is in the printing industry, where one typically wants to maximize the reader's perceived naturalness of a given color image. In broadcasting the goal is to optimize the relationship between image fidelity and frequency bandwidth [5]. If an absolute

truth exists for how the image should look like ("an original"), an often preferred strategy in such research is to develop some weighted measures of "distance to the original", referred to as image metrics. [13]. In medical imaging, on the other hand, there is no original image to compare with. Instead the truth is established for instance by biopsy, and the quality of an image is thus to which extent the decisions based on that image are correct.

One serious attempt to overcome this "distance to the original"-way of defining image quality is due to Janssen [9]. Janssen also regards the quality of an image "not in terms of the visibility of distortions in this image but instead in terms of the *adequacy* of this image as input to the vision stage of the interaction process". He proposes the following two principal requirements that an image of "good" quality should satisfy:

1. the internal representation of the image should be sufficiently precise;

2. the degree of correspondence between the internal representation and knowledge of reality as stored in memory should be high

He goes on to argue that criterion 1 is associated with the usefulness of the image while criterion 2 is related to the naturalness of the image. He then defines image quality as to which extent the image is both useful and natural. With reference to Locke and Descartes it is also quite tempting to associate criterion 1 with objective properties and criterion 2 with subjective ones.

How well Janssen succeeds in disposing of the metric-based assessment of quality is not clear. Engeldrum [7] for instance rhetorically asks how to evaluate art images ("artificial images"), which certainly can be useful, but probably are not natural? It is not always easy to follow Janssen's reasoning, and his ideas have to my knowledge not led to any consensus in the field. What he honorably achieves, however, is to include the purpose/ task of the image (or imaging) into the definition of image quality, and in part he also discusses the observer's role. The question is whether this view can be generalized further in light of Pirsig insight, in which the subjective (observer and task) and the objective (physical) are integrated.

5 A generalized view on the image quality?

In order to answer this question it is interesting to look to statistical decision analysis, which is concerned with decisions in a mathematically formal manner. Often the challenge is to find optimal decision strategies, through calculations of the expected cost / benefit in accordance with the stated purpose (task), technical limitations (physical parameters), and decision strategy (the observer).

In practice, unfortunately, even simple practical decision processes quickly become mathematically unsolvable, and one must resort to idealization. A decision task that proves to be mathematically computable is where an ideal observer (one who is able to utilize all available, including past, information) is to decide whether an exactly defined object on top of a known homogeneous background is present in the image or not. Without entering the mathematical details, what emerges is a quality index ("figure of merit"), a measure of decision quality [8]. For other combinations of observers and tasks the strategy will be different [**?**]. The important lesson here is, however, that this is not related to the quality of the image, but instead to the quality of the *decision*, and which is a combination of the imaging task and observer strategy (subjective) and the fidelity of the imaging system (objective). And it is in this encounter between the subjective and the objective, or between the romantic and the classical as Pirsig would say it, that quality is found.

With this "definition" of picture quality (we choose to retain the term) it might also be interesting to see whether also other types of images than those with purely instrumental purposes, e.g. art paintings, graphics, can be assessed using similar strategies. Some have tried [10, 1], but that is another story.

6 Conclusion

A photographer may deliberately reduce the "technical" quality to achieve other qualities: a striped and / or blurred background can give a sense of motion, unnatural increase in contrast can enhance a dramatic scene, strong emotions can be illustrated by using exaggerated colors, and dreams can be described using the blur and graininess in the image. Even a technically very bad image might thus be valued subjectively good. Moreover, the fact that digital images are easily corrected and manipulated further blurs the distinction between the subjective and the objective.

These everyday examples show that the same image can be evaluated differently depending on the image object and the observer. Will it please or offend, convey a mood, arouse memories, help us sell more cars, give sufficient information to make a correct medical diagnosis, provide identification from a fingerprint, or what else? An artistic illustration of this can be found in Salvador Dali's monumental painting titled "Gala Contemplating the Mediterranean Sea which at twenty feet away is a portrait of Abraham Lincoln" in the Teatro-Museo Dali in Figueras, Spain. If you see it up close you will see only Dali's wife Gala, naked, facing the sea. Going further away, or possibly squinting one's eyes, unveils a portrait of Abraham Lincoln. The question of what is the best image quality thus depends on what you are looking for. With good imaging resolution you see Gala, while bad resolution gives you Lincoln. And alternating between the two creates both an artistic experience and leaves you with a deeper knowledge!

Quality is not a thing. Quality occurs in the meeting point between the subject and object through engagement and caring about: "The difference between a good mechanic and a bad one, like the difference between a good mathematician and a bad one, is precisely this ability to select the good facts from the bad ones on the basis of quality. *He has to care!*" [14]

So next time you are about to take a picture you should think: what is the purpose of this particular picture? What should it be used for - advertising or documentation? Who do you intend to show it to, - your grandchildren on Facebook, or to your insurance company? What do you want to emphasize - Lincoln or Gala? And in the process you might send some thoughts to Pirsig, and remember that quality is about caring about.

References

[1] Álvarez-Gila, A. (2010) *Scene recognition for improved aesthetic quality inference of photographic images.* Master thesis, CIMET, Gjøvik University College

[2] Barfield, O. (1988). *History in English Words.* London: Lindisfarne Press. Opptrykk av 1953 originalutgave.

[3] Barret, H. (1989). *Detection of known signals in inhomogeneous, random backgrounds.* Soc. Photo-Opt. Instrum. Eng. , pp. 176-182.

[4] Behairy, H., & Khorsheed, M. (2005). *Improving Image Quality in Remote Sensing Satellites using Channel Coding.* World Academy of Science, Engineering and Technology. Nr 9

[5] Broderick, T., Harnett, B., Merriam, N., Kapoor, V., & Doarn, C. (2001). *Impact of varying transmission bandwidth on image quality.* Telemed J E HealthSpring , 7 (1), pp. 47-53.

[6] Burgess, A. (1999). *The Rose model revisited.* J. Opt. Soc. Am. A , 16 (3).

[7] Engeldrum, P. (2004). *A Short Image Quality Model Taxonomy.* Journal of Imaging Science and Technology , pp. :160–165.

[8] ICRU (1996). *Medical imaging - the assessment of image quality.* International Commission on Radiation Units and Measurements, ICRU rapport 54,

[9] Janssen, R. (2001). *Computational Image Quality* (PM101 ed.). SPIE Press monograph.

[10] Li, C. & Chen, T. (2009). *Aesthetic Visual Quality Assessment of Paintings*, IEEE Journal of Selected Topics in Signal Processing, vol. 3, no. 2, pp 236-252.

[11] Locke, J. (1689). *An Essay Concerning Human Understanding.* London. Flere nyutgivelser, bl. J.W. Yolton (ed), London 1994.

[12] Nussbaum, P. & Hardeberg J. Y. (2011) *Print Quality Evaluation and Applied Colour Management in Coldset Offset Newspaper Print.* To appear in Color Research and Application

[13] Pedersen, M. & Hardeberg, J. (2009). *Survey of full-reference image quality metrics.* Gjøvik University College report series, 5.

[14] Pirsig, R. (1974). *Zen and the Art of Motorcycle Maintenance. An Inquiry into Value.* New York: Bantam Books.

[15] Pirsig, R. (1992). Lila: an Inquiry into Morals. New York: Bantam Books.

[16] Tapiovaara, M. (2006). *Relationships between Physical Measurements and User Evaluation of Image Quality in Medical Radiology - a Review.* Strålsäkerhetscentralen (STUK) Finland, STUK-rapport A219, Helsinki.

Adaptive Quantization Based on Saliency Maps at Macroblock Level in H.264/AVC

Victor Medina, Fahad Fazal Elahi Guraya, Faouzi Alaya Cheikh
Faculty of Computer Science and Media Technology
Gjøvik University College
Gjøvik, Norway
vjmedina@gmail.com,{fahad.guraya,faouzi.cheikh}@hig.no

Abstract

The rate control algorithm in H.264/AVC determines the amount of bits used to encode each frame based on a Quantization Paramter(QP). This parameter is calculated at run time and aims to achieve a given Mean Average Difference (MAD) between the pixels in the frame. The same QP is used for the entire frame, which can reduce the encoding efficiency of the algorithm. We propose an alternative to overcome this problem based on adjusting the QP at macroblock level. The new QP is computed from a saliency map so that more bits are allocated to more salient pixels. This algorithm adapts to the behaviour of the Human Visual System (HVS), thus obtaining visual results of higher quality while maintaing a high compression rate. Our experiments have produced results of higher quality than those obtained with the standard baseline profile of the JM 18.0 reference encoder, while producing files which are only slightly bigger.

1 Introduction

Rate control is the neccesary mechanism used in video standards like MPEG-2, MPEG-4 and H.264/AVC to control the size of the encoded videos so that they meet the bandwidth requirements of the given video application. In this way, we can be sure that it will be possible to feed data to the applications fluently. Many different approaches have been proposed to control the bitrate of the encoded videos while maximizing their quality. In this model, we propose to use information related to the saliency of the pixels in order to increase the quality of the results at the expense of slightly increasing the resulting file size. The saliency maps are constructed from a Visual saliency model that combines several features, static (like color, intensity, orientation and face) and dynamic (motion), by using an averaging function [2].

Normally, all the pixels are allocated the same amount of resources by the encoder, regardless of their importance to the Human Visual System (HVS)[4] – which usually centers its attention on human faces, disregarding the background [5] – or their possible interest for a human observer in general. For example, an observer might unconsciuosly be more attracted towards areas such as human faces, or pay more attention to familiar objects that might be of interest in specific situations – like, for instance, a suspicious backpack in a surveillance video, or the main character's face in a feature film. In this model, we take those observations into account by assuming that the more salient pixels are those which viewers are more likely to look at, thus encoding them with a higher amount of bits – which, in turn, results in a higher quality; similarly, less salient pixels are encoded with a lower quality. In this model, we try to compensate the higher amount of bits allocated to more salient pixels, by lowering the quality of less salient ones, therefore trying to maintain the mean average difference.

For this experiment we were given the saliency maps before hand, so the calculation of the saliency maps was not actually implemented in the encoder. Nonetheless, there is an important issue to keep in mind when integrating both modules in a working application. Normally, the motion information is needed in order to obtain a proper saliency map; however, the motion

information in H.264/AVC is available only after the Rate Distortion Optimization (RDO) has been performed, for which the value of the QP is needed. A possible solution to this issue is to estimate the actual MAD information from the available spatial and temporal MAD [4] – a more detailed explanation can be found in [7].

2 Computation of the quantization parameter

We propose an alternative model to compute QP. This parameter is computed for every macroblock in each frame from the values of the pixels in the corresponding saliency map. The number of bits allocated for a given frame in the standard H.264/AVC rate control algorithm is computed based on a Mean Average Difference (MAD) criteria. The bitrate is related with the MAD and QP according to the following formula [3]:

$$T_i = c_1 \frac{MAD_i}{Qstep_i} + c_2 \frac{MAD_i}{{Qstep_i}^2} - h_i \tag{1}$$

where T is the assigned bitrate for the basic unit, Qstep is the quantizer step size (from which QP is computed), h_i corresponds to the number of bits due to overhead data, and c_1 and c_2 are model coefficients.

Normally, the MAD is fixed in the encoder configuration, and it is QP that gets updated for every basic unit to maintain a similar MAD for every frame; in our model, however, we set the value for QP from the pixels in the saliency map, and it is the MAD that changes accordingly. This way, we are controlling the quality of the resulting video, whereas the MAD is maintained at macroblock level – as opposed to the standard model which is computed at frame-level.

For each macroblock, the corresponding macroblock is obtained from the saliency map of the frame it belongs to. The intensity of the pixels in the saliency macroblock – in the interval [0-1] – is then averaged and multiplied by the upper bound of QP – see the *Experiment setup* section. This way, we obtain the corresponding final QP for the macroblock, and the same value will be used for the entire macroblock. A new QP will be computed for each new macroblock.

3 Experiment setup

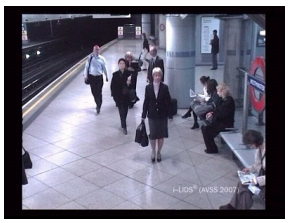


Figure 1: Frame 547 from video 6

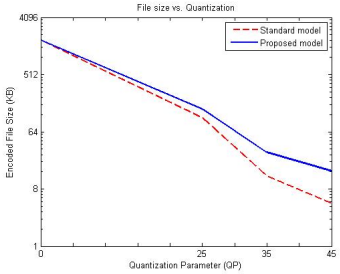
In order to be able to compare our results, we need a reference model. The reference software used for this experiment was JM 18.0 [1]. The tests were run over a set of 6 surveillance video files. Videos 1, 2 and 6 belong to the iLIDS dataset of the IEEE International Conference AVSS

2007, and contain images from a surveillance camera at a subway station. Video 3 shows a man picking an object from a store and leaving. Video 4 shows people passing by at a waiting room in a train station[1][2].

Several scenarios were tested in order to find the one that offered the best compromise between quality improvement and file size decrease. These scenarios were all tested on video 6 (see Fig. 1), both for the reference software and our proposed model, chosing the parameters in a way that they produce similar differences for both models. This is important because, otherwise, when we compare the metrics later on, we cannot be sure of whether the different results are due to the model itself or to the parameters we chose.

Firstly, an upper bound was set experimentally for QP; this upper bound will be used for the most salient pixels. The upper bound for the quantization parameter (QP) was set to 25 by studying the influence of several QP values on the quality and size of the resulting video. The chosen QP corresponds to half of the maximum allowed quantization (51), and proved to be an upper bound after which some pixelation starts to notice in the frames. Similarly, the size of the resulting file for a sample video fragment from video 6 was measured for both models. In figure 2, we can see that the resulting file size decreases greatly when QP reaches 25, to then continue decreasing more slowly after 35.

Once the interval for QP was set, the next step is to find the configuration profile to be used in the reference software. This profile will determine the value of several other encoding parameters like the number of B frames, or the size of the group of pictures (GOP). For this experiment, the baseline profile was used and, therefore, no B frames were used. Some tests were carried out in order to determine whether or not using B frames would add a substantial improvement to the results, but the additional encoding time was higher than the improved performance.



QP	File size Standard model (KB)	File size Proposed model (KB)
0	4032	4032
25	176	250
35	16.8	43.5
45	5.66	20.6

Figure 2: Effect of using different QP values on the file size of the video. The values were computed for the first 25 frames of video 6.

4 Results

We will now show the experimental results obtained with our proposed model, and compare them with those obtained for the standard H.264/AVC rate control algorithm. The comparison has been both quantitative and qualitative, paying special attention to metrics based on correlation to compare the visual properties of the resulting videos.

The first quality measure used is the Peak Signal to Noise Ratio (PSNR), which objectively measures the mean squared error (MSE) between the original and encoded video frames. The

[1]Video 5 was not used in this experiment.

[2]Keep in mind that the first 51 frames of each video were not used. The frame numbers hereby used suppose that frame 52 is the first frame and one should therefore add 51 to the frame number in order to obtain the corresponding absolute frame number.

higher this metric is, the better the quality of the encoded video will be. However, this metric has many limitations, one of them being that it does not correlate well with subjective video quality measures with human observers [3], because it measures the pixels linearly without considering the properties of the HVS.

As one can see from fig. 5, some results obtained with our model have a PSNR up to 6 dB higher than those obtained with normal quantization. This shows that using saliency maps results in higher quality videos. On the other hand, the resulting videos are slightly bigger – which also means an increment in the bitrate – although the reduction is still considerable.

Another important objective metric is the Structural Similarity (SSIM index)[8]. This index represents how similar two images are, in a scale from 0 to 1, where 1 means that the images are identical. Unlike the PSNR, this metric takes into account the visual properties of the images and, therefore, provides results more related to a subjective observation. When observing the encoded video frames we can see that, despite having reduced the bitrate and file size so much, there are no visual differences between the encoded frames and the original ones. This is consistent with the results obtained for the SSIM indexes, which are very close to 1. Fig. 3 shows an example of the improved quality for video 1; we can see that the SSIM disparity map is much closer than the one obtained with the standard quantization – warmer colors indicate that the pixels are close to the original, whereas cold colors indicate very distant pixels.

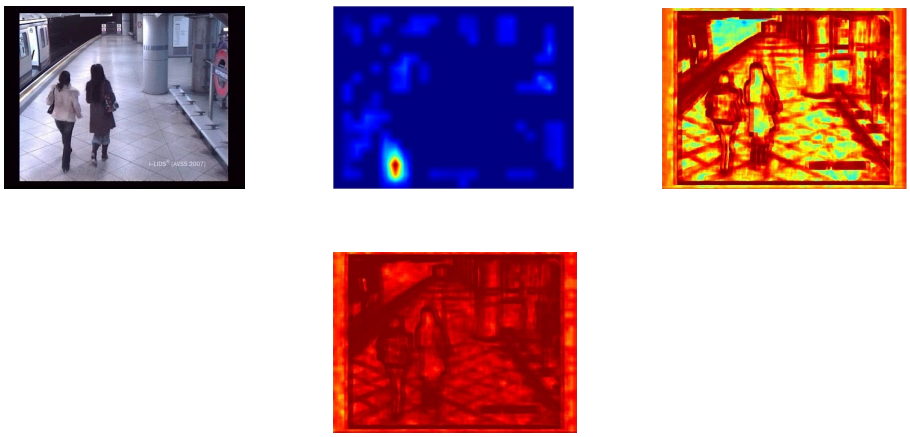


Figure 3: Left to right: frame 158 of video 1; saliency map for the frame; SSIM disparity map obtained with the standard model; SSIM disparity map obtained with our model.

The Root Mean Square Error (RMSE or residual image) was also obtained to study how different the encoded frames are from the original. It is desired to obtain values as low as possible for this metric, because that means that the video has not suffered a big change from the original. Fig. 5 shows that the videos encoded with our model are always closer to the original video and, sometimes, more than a 110% closer.

Other correlation metrics used to compare the results are the Sum of Squared Distortion (SSD), Sum of Absolute Differences (SAD) and Sum of Hamming Distances (SHD) [9]. The tables in fig. 5 show the results obtained for all these metrics. Fig. 4 shows how our model produces videos with a higher visual correlation with the original, which is what we expect from this model.

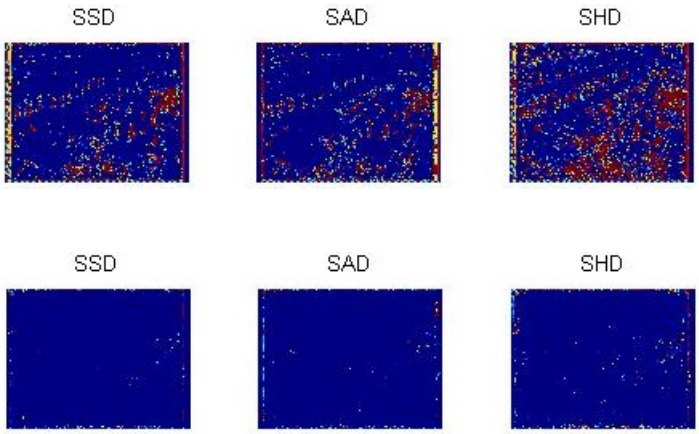


Figure 4: Disparity maps obtained for the SSD, SAD and SHD for frame 121 of video 4, after encoding with the standard quantization model (Top row) and using saliency maps (Bottom row)

	JM 18.0						
Video	Average PSNR (dB)	SSIM index	SSD	SAD	SHD	RMSE	Frame
1	42.5508	0.9404	0.0397	0.0379	0.0643	2.2172	158
2	42.5332	0.9419	0.0376	0.0375	0.0631	2.2701	704
3	41.4548	0.9653	0.0327	0.0330	0.0472	2.4887	210
4	41.7478	0.9480	0.0291	0.0284	0.0588	2.4486	121
6	42.4580	0.9423	0.0369	0.0444	0.0660	2.2778	547

	Proposed Model						
Video	Average PSNR (dB)	SSIM index	SSD	SAD	SHD	RMSE	Frame
1	44.6355	0.9851	0.0113	0.0120	0.0144	0.9484	158
2	43.1027	0.9481	0.0336	0.0345	0.0540	2.2107	704
3	44.3521	0.9885	0.0084	0.0056	0.0098	0.9922	210
4	47.8359	0.9891	0.0021	0.0035	0.0044	1.0818	121
6	48.0546	0.9870	0.0044	0.0059	0.0073	1.0033	547

Figure 5: Metric results obtained for one sample frame of each of the video sequences for both models. The PSNR was computed for the entire videos.

	Original		JM 18.0		Proposed Model	
Video	Encoded Filesize (KB)	Bitrate (Kbps)	Encoded Filesize (KB)	Bitrate (Kbps)	Encoded Filesize (KB)	Bitrate (Kbps)
1	138829	34546.94	6997	1741.07	18290	4551.38
2	203161	34701.70	10831	1849.94	16517	2821.23
3	60394	29379.01	3583	1742.75	8946	4351.50
4	189848	32509.05	9237	1581.58	49106	8408.76
6	253126	35801.18	14766	2088.37	66352	9384.53

	JM 18.0		Proposed Model	
Video	Filesize reduction (%)	Bitrate reduction (%)	Filesize reduction (%)	Bitrate reduction (%)
1	94.96	94.96	86.83	86.83
2	94.67	94.67	91.87	91.87
3	94.07	94.07	85.19	85.19
4	95.13	95.13	74.13	74.13
6	94.17	94.17	73.79	73.79

Figure 6: File size and bitrate reduction achieved by both models for the entire encoded video sequences.

5 Conclusion

Observing the results in fig. 5, we can safely conclude that using saliency maps to encode the macroblocks produces better results than using the standard rate control algorithm. As we already mentioned before, this is due to the fact that the saliency maps contain information about which elements (pixels) in the scene might be more relevant to the observers and, therefore, allows the encoder to allocate a higher amount of bits and resources to those pixels, disregarding less relevant information –by encoding it with a lower quality. This, however, should be transparent to the observers because they will – unconsciously – not pay much attention to those pixels with lower quality.

This model could be a good solution for applications where bitrate and encoding time are not crucial. We have only tested with surveillance videos, but the model can be easily extrapolated to any other type of videos where non-salient information is not to be paid as much attention by the observers.

In general, the saliency maps are not available to the encoder beforehand, so a practical application of this model should integrate in the encoder the algorithm to generate the maps.

All the results above have shown that this approach produces better quality results than the standard model included in the JM 18.0 reference software. Nonetheless, they also show that our model tends to increase slightly the size of the encoded videos and, consequently, their bitrate [3], so one must pay attention to whether or not this approach could be adequate for the desired application, for it may otherwise affect its running time and reduce its performance. In some cases, it might even not be a feasible solution if, for instance, it were to used in ral-time communications, or other similar applications where encoding time and bitrate are crucial.

[3]We have not done any studies about the encoding time but, since using saliency maps in the encoder requires several additional operations, this will most likely incurr in added computation time.

References

[1] JM 18.0 - http://iphome.hhi.de/suehring/tml/download/

[2] F.F.E. Guraya, F.A. Cheikh, A. Tremeau, Y. Tong, H. Konik, *"Predictive Saliency Maps for Surveillance Videos"*, Proceedings of the 2010 Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science, IEEE Computer Society, DCABES '10, pp. 508-513, 2010.

[3] Iain E. Richardson *"The H.264 advanced video compression standard"*. John Wiley & Sons, Ltd, 2nd Edition, 2010.

[4] Y. Liu, Z. Li, Y.C. Soh, *"Region-of-interest based resource allocation for conversational video communication of H.264/AVC"*. IEEE Transactions on Circuits and System for Video Technology, Vol.18, no.1, January 2008, pp.134-139.

[5] K.C. Lai, S.C. Wang, D. Lun, *"A rate control algorithm using human visual system for video conferencing systems"*. International Conference on Signal Processing, Vol.1, August 2002, pp.656-659.

[6] L. Tong, K.R. Rao, *"Region-of-interest based rate control for low-bit-rate video conferencing"*. Journal of Electronic Imaging, Vol.15, no.3, July 2006.

[7] Y. Liu, Z. Li, Y.C. Soh, *"A Novel Rate Control Scheme for Low Delay Video Communication of H.264/AVC Standard"*. IEEE Transactions on Circuits and System for Video Technology, Vol.17, no.1, January 2007, pp.68-78.

[8] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, *"Image quality Assesment: From Error Measurement to Structural Similarity"*. IEEE Transactions on Image Processing, Vol.13, no.1, January 2004.

[9] S. Ahuja *"Correlation based similarity measures"*. http://siddhantahuja.wordpress.com/2010/04/11/correlation-based-similarity-measures-summary/ April 2010.

How Many Pixels Does it Take to Make a Good $4'' \times 6''$ Print?

Pixel Count Wars Revisited

Michael A. Kriss
MAK Consultants
506 NE 193RD Avenue
Camas, WA 98607, USA
makriss@comcast.net

Abstract

In the early 1980's the future of conventional silver-halide photographic systems was of great concern due to the potential introduction of electronic imaging systems then typified by the Sony Mavica analog electronic camera. The focus was on the quality of film-based systems as expressed in the number of equivalent number pixels and bits-per-pixel, and how many pixels would be required to create an equivalent quality image from a digital camera. It was found that 35-mm frames, for ISO 100 color negative film, contained equivalent pixels of 12 microns for a total of 18 million pixels per frame (6 million pixels per layer) with about 6 bits of information per pixel; the introduction of new emulsion technology, tabular AgX grains, increased the value to 8 bit per pixel. Higher ISO speed films had larger equivalent pixels, fewer pixels per frame, but retained the 8 bits per pixel. Further work found that a high quality $3.5'' \times 5.25''$ print could be obtained from a three layer system containing 1300 x 1950 pixels per layer or about 7.6 million pixels in all. In short, it became clear that when a digital camera contained about 6 million pixels (in a single layer using a color filter array and appropriate image processing) that digital systems would challenge and replace conventional film-based system for the consumer market. By 2005 this became the reality. Since 2005 there has been a "pixel war" raging amongst digital camera makers. The question arises about just how many pixels are required and are all pixels equal? This paper will provide a practical look at how many pixels are needed for a good print based on the form factor of the sensor (sensor size) and the effective optical modulation transfer function (optical spread function) of the camera lens. Is it better to have 16 million, 5.7-micron pixels or 6 million 7.8-micron pixels? How does intrinsic (no electronic boost) ISO speed and exposure latitude vary with pixel size? A systematic review of these issues will be provided within the context of image quality and ISO speed models developed over the last 15 years.

1 Historical Introduction

As early as 1969 Kodak Harrow Research Laboratories were actively building a 1000 line, electronic CRT based scanner and printer. The goal was to scan color transparencies and in an analog fashion apply the electronic equivalent to an un-sharp mask and then using a CRT based printer create a very high quality color print. The unsharp mask was designed in enhance the image on the print between a frequency range of 2 c/mm to 4 c/mm since experiments [10] at the Kodak Research Laboratory in Rochester, NY had indicated that this was the optimum range for visual enhancement of prints viewed at a normal viewing distance of four picture heights. This work was led by Bob Hunt in Harrow and marks the first meaningful move to electronic imaging at Kodak.

In the late 1970's and early 1980's Kodak developed and introduced the Disc Camera system that used circular negatives with 15, 8 mm by 11 mm negative images. It soon became clear the overall quality of the Disc system was below standard. Using the Cascaded Modulation Transfer

Acutance (CMT Acutance) sharpness quality metric developed by Tim Crane [4][7] and later modified by the author [7][8] it was found that the Disc Image was in the "acceptable" image quality range, but not in the "good" or "excellent" quality range found in either the 110 camera systems or the growing compact 35mm camera systems. In order to enhance the quality of Disc images, the Kodak Research Laboratories (KRL) undertook a program to scan the images with a laser (CCD scanners were not well developed at that time), process the images automatically for improved sharpness, less grain (noise) and better tone scale rendition. The noise reduction and enhancement algorithms developed by Bryce Beyer (KRL) and Phil Powell (Harrow) [3] used a moving Hadamard Transform. The tone scale improvements developed by James Alkofer (KRL) used methods based on the Bayesian Statistics of the oil painting through the centuries by the great masters. Based on the algorithms (implemented in general purpose computers or in dedicated, high speed equipment) the CMT Acutance of the images rose from 85 (barely good) to 90, the high end of a "good" rating. In addition, the Alkofer algorithms enabled to compress the tone scale on the negative to the tonal range of the print, creating a much more natural looking print. All this formed the basis for how may "pixels" were required for a "good" print. Using complete models of the Disc film, and scanning and printing systems it was found that a 1300×1788 array would be need for a Disc negative or 1300×1950 for a 35 mm negative. Since there are three layers in each negative, this translates to roughly a 7 mega-pixel image for Disc or a 7.6 mega-pixel system for a 35 mm print. At the time it was estimated that a 35 mm frame contained 18 million pixels at 8 bits per pixel.

Today, with the exception of the Foveon sensor based digital cameras, consumer and pro-consumer cameras contain a single sensor (CCD or CMOS) with pixels counts ranging from about 5 million in the lower end digital cameras to well over 20 million pixels in full-frame 35mm format sensors. So the question that must be addressed is "that if a film image only required about 7 million pixels (using three layers) how many pixels are required in a single sensor camera using a color filter array to encode the color?" This question will be answered in what follows.

2 Image Quality Factors

Image quality is generally defined as some non-linear combination (human perception) of color reproduction, sharpness, noise, exposure latitude and dynamic range. Camera or sensor speed (sensitivity to light) while not a quality factor per se influences the final quality in that as the ISO speed of the camera is increased by means of electronic gain there is a general increase in noise. In this paper color reproduction will not be considered and the focus will be on sharpness, noise as it affects ISO speed, exposure latitude and dynamic range of the output. The single most important parameter will be pixel size as it impacts on system sharpness, ISO speed, dynamic range and exposure latitude.

3 Sharpness

System sharpness is measured by CMT Acutance [7] that is defined by

$$CMT = 100 + 66Log(R_{system}/R_{eye}) \tag{1}$$

where R_{system} is the area under the system MTF curve and R_{eye} is the area under the eye MTF curve. The system MTF includes the MTF of the lens, the MTF of the anti-aliasing low pass optical filter, the geometrical MTF of the pixel, the MTF due to interpolating the sampled

image, the digital enhancement MTF applied to the interpolated image, the laser printing beam MTF and the paper MTF. For these discussions a photographic laser printer will be assumed for optimum quality. The eye MTF is also included at a given viewing distance, normally four viewing heights of the image (for a 4-inch print, that would be 16 inches). The ratio of R_{system}/R_{eye} is normally less than unity unless very strong digital enhancement is applied; hence CMT values of less than 100 are typical. For reference, a very good 35 mm SLR film system using a high quality color negative film would have a $CMT = 96$ for a 4-inch by 6-inch print viewed from 16 inches. Table 1 shows the relationship between CMT Acutance values and perceived quality [8].

CMT Range	Quality
92 and Above	Excellent
86 to 91	Good
80 to 85	Acceptable
76 to 80	Poor
75 and Below	Unacceptable

Table 1: CMT values versus quality.

In today's digital still cameras, the biggest image sharpness factor is the quality of the lens. To understand the relationship between the lens and the pixel size consider a set of ideal diffraction limited lens that depend only on the F-number of the lens and the wavelength of light to define the MTF of the lens or it counterpart, the point spread function (the Airy Function) in the focal plane of the sensor [5]. Table 2 shows the diameter of the Airy Function as a function of the F-number for green light of 550 nanometers of a diffraction limited lens along with the cutoff frequency in c/mm in the plane of the sensor. As can be seen from Table 2, ideal lenses have optical spread functions that will more than likely have spread functions that exceed the size of the pixels, thus the camera systems are dominated by the quality of the lens.

F/Number	Airy Diameter (microns)	Cutoff Frequency (c/mm)
1.4	1.88	1300
2.8	3.76	484
4.0	5.37	339
5.6	7.52	242
8.0	10.74	162
11.0	14.76	123
16.0	21.47	85

Table 2: Airy diameter and cutoff frequency as a function of F-Number

Survey of high quality 35 mm primary lenses (one focal length) show that the frequency values at 50% MTF translate to diffraction limited lenses ranging from F/8 to F/23. These are expensive lenses. Data for the MTF of these lenses used in compact digital cameras is not generally available. Consider the Nikon CoolPix S51 compact digital camera. The sensor has an array of 2428×3237 pixels, each 1.75 microns square. Photographic experiments using a test target with 48 line pairs per inch indicated the following. When the camera was used at both the 6.3 mm focal length setting and the 18.9 mm focal length setting at various "shooting distances" aliasing was noticed (in the color and black and white modes) that agreed with a sampling frequency of 286 c/mm and a Nyquist frequency of 143 c/mm. This sampling frequency indicates

the use of a Bayer type CFA. In the black-and-white mode, a spatial frequency on the sensor of 443 c/mm (focal length of 6.3 mm at a distance of 1.438 meters) showed aliasing and very low contrast. This and other results infer that the lens quality in the camera is between diffraction-limited lenses of F/8 and F/5.6. This also means that the Airy Function diameter falls between 10.7 microns and 7.5 microns, both much larger than the 1.75 micron pixel size. Thus even these "good" lenses do not make full use of the pixel resolution. A similar study of a less expensive Nikon CoolPix S220 with similar optics to the S 51 indicated an equivalent F/11 diffraction limited lens. In what follows, it will be assumed that a typical higher end compact camera will have a diffraction limited lens of F/8 and that of a typical digital SLR camera will have a diffraction limited lens of about F/11. At F/11 the Airy Function diameter is 14.8 microns, again larger than the typical pixel size in DSLR sensors, which range from 4 microns to 8 microns.

Based on the color aliasing from the 48 lp/inch test target, there must be an optical pre-filter in CoolPix S51, CoolPix S220 and the DSLRs studied, thus in all cases it will be assumed that the optical pre-filter will shift the image one whole pixel for optimum results.

The following analytical simulations of digital cameras systems is defined by the following imaging chain:

- Camera lens (F/8 MTF for compact cameras and F/11 MTF for DSLRs)
- Optical Pre-filter (one pixels shift MTF)
- Image Sensor (the geometric MTF is used, but electron diffusion within the sensors will lead to lower MTFs)
- Bayer CFA and interpolation MTF
- Enhancement MTF (varies to ensure that image is enhanced in an optimum fashion for given viewing distance,
- Laser beam MTF (Gaussian) designed not to show line structure on the print.
- Photographic color paper MTF
- Eye MTF based on viewing distance

These eight MTF functions are cascaded (allowing for magnification) in the viewing plane of the photographic paper and they are used in Equation [**?**] to obtain the CMT Acutance value. The compact digital cameras were modeled with pixel sizes ranging from 1.5 microns to 4 microns, while the DSLRs were modeled with pixel sizes ranging from 4 microns to 12 microns. The compact digital cameras were assigned formats of 6×4.5 mm, 7.2×5.4 mm and 8×6 mm. The DSLRs were assigned formats 12×18 mm (roughly DX format) and 24×36 mm (roughly FX format). Overall, the CMT values varied in a roughly logarithmic manner when plotted versus effective lines per picture height or total pixels; see Figure 1. The model used is one-dimensional and thus the plot of CMT values versus the number of lines per picture height better reflects the logarithmic nature of the quality improvement with the increase of pixel count. The results clearly indicate that the compact digital cameras are less sharp than those of the DX or FX formats for the DSLRs. When the total pixel count exceeds 6.5 million the quality of the images for the standard 4-inch by 6-inch print viewed from 16 inches (four viewing heights) should be "good" or better. Those below will be acceptable. For all cases when the pixel counts are equal for cameras, but one is based on a larger format and larger

pixels, the larger format camera will have about a two CMT value improvement. A one CMT unit improvement can be seen. Thus, while the total number of lines or pixels has the strongest impact on sharpness quality, larger formats at equal pixel counts give improved sharpness. This is true for the DX and FX formats where all systems show "good" or "excellent" quality. The impact of format size is very evident when one has a 3000-line image in the DX and FX systems; the CMT jumps from about 93 to 95, both "excellent". This improvement in sharpness quality will be amplified when we consider the gain in speed and exposure latitude when the pixel size increases from 6 microns to 12 microns.

In summary, as the pixel count and format area increases; there are steady increases in sharpness quality. The first "excellent" image in the compact formats (as defined above) is for a 1.5-micron pixel in an $8mm \times 6$ mm format for a 21 mega-pixel camera (4000 lines). For the DX format the first "excellent" image is a 6 micron pixel, 3000 lines, 13.5 mega-pixel system. For the FX format, the first "excellent" image is a 10 micron, 2400 lines, 8.64 mega-pixel system. Figure 1 shows the CMT values as a function of pixel count for all sensor sizes and pixel sizes.

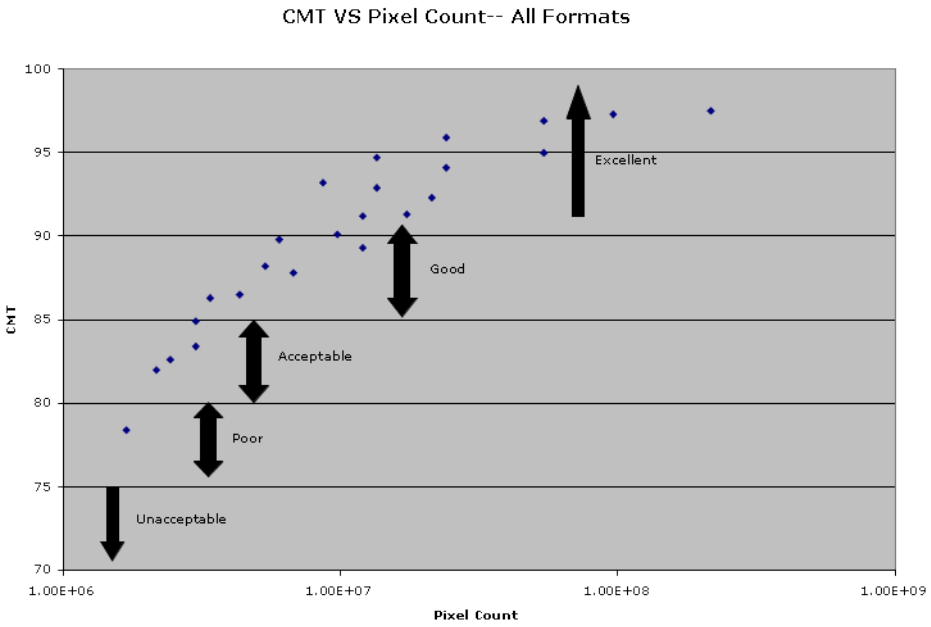


Figure 1: CMT versus pixel count.

4 Sensor Speed, Exposure Latitude and Dynamic Range

As seen from above, the sharpness quality of a digital camera depends on a combination of the pixel size and size of the sensor (along with other factors) where, as the sensor size increases and pixel dimensions decrease, the sharpness improves (more pixels). The effective speed, exposure latitude and dynamic range depends only on the size of the pixel in terms of light gathering ability, photo-electron generation and storage capacity and the noise electrons generated in the sensor and when the signal is read out and amplified. The important factors are listed below:

- Micro lenses
- Color Filter Arrays (CFA)
- Effective Pixel Size or Fill Factor
- The Quantum Efficiency of the sensor
- The effective storage area for charges in the pixel and shift registers (for CCDs)
- Dark Current
- Fixed Pattern Noise
- Shot Noise
- Readout noise

For the purpose of this paper it will be assumed that the micro lenses have effectively increased the fill factor for each pixel to unity and that the peak quantum efficiency for all sensors will be 50% at 700 nanometers and that the green element of the CFA has 50% transmittance at 550 nanometers. Both IR and UV filters are used to limit the generation of photo-electrons to the visible region of the electromagnetic spectrum. The noise levels assume Double Correlated Sampling at the read-out amplifier and that the total "base" noise will be limited to about 13 electrons RMS. Based on a search of the literature [1] a value of 3×10^{10} electrons per cm^2 per volt will be assumed for the electron density possible for all sensors and this value is based on the full pixel area; in reality the electron density is higher, but no sensor architecture allows for a storage area equal to the pixel area and the above number is an average over several types of sensors when the full pixel area is assumed active. For compact digital cameras a pixel gate voltage of 3.7 volts will be assumed and for DSLR cameras a gate voltage of 7.3 volts will be assumed (equal to the battery voltages used in the cameras). Also, all sensors are assumed to be front illuminated and do not include the more efficient back illuminated sensors now appearing on the market [12].

The speed defined in this paper will refer to the "intrinsic" speed of the sensor and is independent of any amplification within the camera to achieve higher speeds. The speed evaluation will be based on a model developed by the author [9] and not the ISO Standard 12232 [6]. The reason for using the speed determination developed by the author is that it lends itself to the physics of the sensor and not the final digital output of the camera that is incorporated into the ISO standard. The speed, S, is defined by

$$S = 0.8/E_{s/n} = 10 \tag{2}$$

where $E_{s/n}$ is the exposure in lux-seconds in the plane of the sensor that gives a signal-to-noise ratio of 10. The formula treats the sensor in the same manner as conventional black-and-white negative material [2] and thus fits into standard camera metering systems and also follows the rule of thumb that an excellent image would be captured at an exposure at f/16 at a shutter speed of 1/S. The exposure latitude will be expressed in the ratio of scene luminance that can be recorded and is based on the full well capacity of the pixel and the base noise. The dynamic range is defined as $20Log10[maximum - signal/base - noise]$ in dBs. Table 3 summarizes the calculations.

The results indicate a clear linear dependence of speed and full well capacity the area of the pixel. For a 2 micron pixel a speed of 15 is calculated. Most digital cameras with this size

pixel quote an ISO speed of about 50 to 100. Keep in mind that these stated speeds are based on the ISO Standard and will include any amplification done within the camera prior to the D/A converter. What is important is the rapid increase in speed, full well capacity, exposure latitude and dynamic range as a function of pixel area. Clearly larger pixels are better. A study of scene luminance [11] has shown that the average scene luminance ratio is about 160:1 and that that the 90% cumulative value is about 400:1 in exposure latitude. Thus a 4-micron pixel (at 3.7 volts) would be required to capture the mean scene luminance ratio accurately and a 5-micron pixel (at 7.3 volts) would be required to capture most scenes if the camera metering system is accurate. Clearly increasing the Quantum Efficiency, QE, and green transmittance of the CFA would improve these results. For example, increasing the peek QE to 70% and the green transmittance to 80% gives a speed of 33 and exposure latitude of 45:1. The exposure latitude does not change since the pixel size has remained constant, but the increased sensitivity to light has increased the speed from 15 to 33. If the peak QE was 100% and the green peak transmittance was also 100%, the 2-micron pixel speed would increase to 59.

5 Exposure Accuracy and Exposure Consistency

Seven digital cameras (Nikon D700, D70, S5, S51, S220, S4000, S8000) and one film camera (Nikon N90) were used to check the exposure accuracy and resulting image exposure consistency. The goal was to see if all cameras gave a consistent exposure value on a common "real" target. Each camera was used in its automatic mode and the various lens and zoom combinations were adjusted to cover the same scene content. Figure 2 shows the scene used for this experiment.

Figure 2: Scene used to establish accuracy and consistency of exposure system.

The EXIF image files created by all the digital cameras contain the ISO speed used by the camera, the exposure time and the f-number of the lens. The film camera was set at ISO 400 and the exposure time and f-number of the lens was noted from the viewfinder. Equation [**?**] gives the exposure on the sensor (or film) in terms of the lens f-number, F/N, and exposure time, T, and Equation [**?**] gives an exposure index based on the exposure and the ISO speed of the sensor (film), S. The exposure index should be constant for all cameras if they all are calibrated in the same fashion.

$$Exposure \propto T/(F/N)^2 \tag{3}$$

$$ExposureIndex = T \times S/(F/N)^2 \tag{4}$$

The test was run on the same scene on a bright day and an overcast day to gather information on a wide range of scene luminance. The exposure index varied from 0.0057 to 0.0092 for the bright day with an average value of 0.0077 and a standard deviation of 0.0016. For the overcast day the exposure index varied from 0.0048 to 0.0072 with an average value of 0.0063 and a standard deviation of 0.0009. Given that the automatic exposure mechanisms are not continuous it would appear that the digital cameras have a consistent exposure algorithm. The film results gave different exposure index values for bright and cloudy days, all higher than the digital camera values, indicating that the film camera may have a more complex or different exposure system than the digital cameras. Given the above it would appear that the digital images should all look similar for exposure and tone scale. However the actual images show considerable variation in overall tone scale. The images from the D700 and D70, with their larger pixels, produce images with lower apparent contrast and are able to record more information in the deep shadows than the compact digital cameras with smaller pixels, and less exposure latitude. Since all the output files from the seven digital cameras are 8-bits deep, those with lower exposure latitude will tend to have higher contrast in the final image.

Figure 3 shows a plot of exposure versus camera image file code value (0-255) for six model digital cameras (based on pixel sizes from 1 micron to 8 microns). Note that the compact cameras, with their smaller pixels, all have a higher contrast than the two SLR type digital cameras, with their larger pixels.

Using pictures taken by the D700 or D70 SLR digital cameras it was possible to "create" approximations of the images that came from the compact cameras by using Photoshop® and reducing the output range by the number of code values indicated by difference in exposure latitude of the between the larger pixels and smaller pixels. It seemed that the cameras tend to expose for the brightest area of the image.

6 Conclusions

It has been shown, in a consistent set of models, the relationship between image quality, speed and exposure latitude. This enables one to choose a digital camera based on the required use. For compact cameras consider Table 3 below.

If the goal is to have a good to very good picture in terms of sharpness, one could use the $CMT = 90$ pixels given in Table 4. Using a 2 micron pixel would, based on Table 3, provide exposure latitude of 45:1, would not cover the average luminance ratio of 160:1. Moving to the 2.5 micron pixel would provide exposure latitude of 71:1, still well below the average luminance ratio. If one would settle for a good image, $CMT = 86$, then using a 3.6 micron pixel would provide an exposure latitude of 141:1, close to the average of 160:1. Hence in small, compact digital cameras, while it is possible to get sharp images (at relatively low speeds), it will be difficult to obtain wide exposure latitude.

If one wishes to use the DX format DSLR, then the models indicate that either a 5 micron (8.64 mega-pixel) or a 4 micron (12 mega-pixel) camera will provide very good images and cover exposure latitudes in excess of 354:1.

If one wishes to use a FX format DSLR, then a 10 micron (8.64 mega-pixel) camera will produce excellent images with virtually unlimited exposure latitude. In the FX format, going to smaller pixels will provide increases in sharpness with only a small cost in exposure latitude.

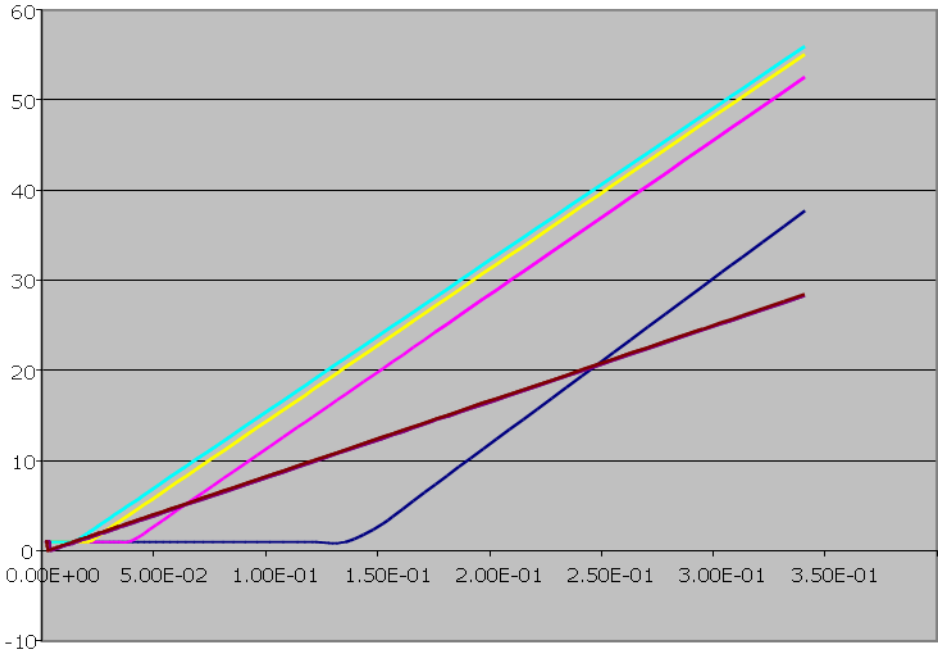


Figure 3: Image code values versus exposure for a series of digital cameras with pixels ranging from 1-micron to 8 microns, The plots to the left are for the cameras with 2, 3 and 4 micron pixels and the one to the right is for a 1 micron pixel where the slower speed point is clearly evident. The lower sloped plot is for both the 6 and 8-micron pixels.

Pixel Size, microns	Pixel Area, sq-microns	Speed for green pixels	Full Well Capacity, electrons	Exposure Latitude Ratio	Dynamic Range, dB
1	1	4	1,110	11:1	39
2	4	15	4,440	45:1	51
3	9	33	9,990	100:1	58
4	16	59	17,800	178:1	63
5	25	83	54,800	562:1	72
6	36	118	78,800	794:1	76
7	49	166	107,000	1122:1	78
8	64	209	140,000	1413:1	81
9	81	263	177,000	>1413:1	83
10	100	295	219,000	>1413:1	85

Table 3: Summary of speed, full well capacity, exposure latitude and dynamic range as a function of pixel dimensions.

An advanced amateur or professional might prefer an 8 micron pixel (13.5 mega-pixel) camera that gives excellent sharpness and the potential for over 1000:1 exposure latitude.

Given the above, it is not clear why one would, in the FX format, use pixels less than 8 microns since the gains in sharpness are small.

The impact on enlargements depends on the viewing distance. If one enlarges an image and

		Sensor Size		
		$6mm \times 4.5mm$	$7.2mm \times 5.4mm$	$8mm \times 6mm$
CMT	86	2.5	3.2	3.6
	90	1.5	2	2.5
	93	0.8	0.9	1

Table 4: Pixel size in microns to realize a given CMT Acutance value for three sensor formats.

still views it from "four viewing heights", the image quality will be preserved if not increased due to the effectively higher MTF response of the paper. On the other hand, if one makes a 2x enlargement and views the final print at the same distance, say 16 inches, the sharpness will fall off considerably. In this case, using smaller pixels will preserve the sharpness of enlargements viewed at close distances and with no loss of sharpness at normal print sizes (it will be hard to see the increased sharpness in the normal viewing distances), but at some cost to the exposure latitude due to the smaller pixels.

When such cameras are used in the 14-18 bits Raw format, subsequent image processing can produce outstanding images. Note that JPEG compression will limit one to 8 bits per color and thus a good deal of the exposure latitude will be lost. The best compression ratio in most digital cameras for JPEG compressed images is about 4:1. Errorless TIFF compression can be used to preserve quality. However, it was noted that al of the compact cameras showed JPEG blocking artifacts at higher print magnifications. The EXIF file information indicated that the compact cameras used higher compression ratios than the larger cameras.

Table 5 summarizes the size of pixel and pixel count for a wide range of photographic uses based on the results of this study.

References

[1] The author gathered a lot of CCD and CMOS sensor data on full well capacity and found an average value for the electron density based on the full pixel area and not just on the active storage area, which is difficult to determine.

[2] J. Altman. Densitometry of black-and-white materials. In Thomas Howard James, editor, *Theory of the Photographic Process*, chapter 17, pages 501–509. Macmillan Publishing Co, Inc., New York, USA, fourth edition, December 1977.

[3] B. Bayer and P. Powell. A method for the digital enhancement of unsharp, grainy photographic images. In T. S. Huang, editor, *Advances in Computer Vision and Image Processing*, volume 2, chapter 2. JAI Press Inc., 1986.

[4] E. Crane. An objective method of rating picture sharpness: Smt acutance. *J. SMPTE*, 73:643–647, 1964.

[5] E. Hecht and A. Zajac. *Optics*. Addison-Wesley, Reading, UK, 1979.

[6] ISO. *ISO 12232: Photography - Electronic Still Picture Cameras: Determination of ISO Speed*. International Organization for Standardization, Geneva, Switzerland, 1997.

[7] M. Kriss. Image structure. In Thomas Howard James, editor, *Theory of the Photographic Process*, chapter 3, pages 628–630. Macmillan Publishing Co, Inc., New York, USA, fourth edition, December 1977.

[8] M. Kriss. Image analysis of discrete and continuous systems: Film and ccd sensors. In *Proc. SPIE*, volume 1398, pages 4–14, 1990.

[9] M. Kriss. A model for equivalent iso ccd camera speed. In *Digital Solid State Cameras: Designs and Applications, Proc. SPIE*, volume 3302, pages 56–57, 1998.

Format	Use	Pixel Size micron	Mega-pixels	CMT	Green Intrinsic Speed	Lati-tude	Comments
$4 \times 6mm$	General	2	6	86	15	45:1	Good sharpness, poor latitude, good for LCD monitor
$6 \times 9mm$	General	3	6	89	33	100:1	Good sharpness and reasonable latitude
$8 \times 12mm$	General	3.5	7.8	92	42	141:1	Excellent sharpness, will capture most normal scene illumina-tions
$12 \times 18mm$	Advanced	5	8.6	93	83	562:1	Excellent camera for advanced consumer
$24 \times 36mm$	Advanced	10	8.6	94	295	$>$ 1412:1	Excellent camera for advanced consumer
$24 \times 36mm$	Profession-al	6	24	97	118	794:1	Ideal for all non-sports photography
$24 \times 36mm$	Studio	5	24	98	83	562:1	Ideal for any studio shots where lighting and exposure time can be controlled

Table 5: This table summarizes the projected design of digital cameras based on use along with the relative intrinsic speed, sharpness (in terms of CMT values) and exposure latitude.

[10] M. Kriss, R. Mickelson, and N. Nail. Image structure and visual sharpness. 1971.

[11] C. Nelson. Tone and color reproduction. In Thomas Howard James, editor, *Theory of the Photographic Process*, chapter 19, pages 547–549. Macmillan Publishing Co, Inc., New York, USA, fourth edition, December 1977.

[12] Sony. `http://www.sony.net/Products/SC-HP/cx_news/vol59/np_imx050cqk.html`.

Biography

Dr. Kriss received his BA, MS and PhD in Physics from the University of California at Los Angeles. He joined the Eastman Kodak Research Laboratories in 1969 and until his retirement in 1993 worked on both conventional and digital imaging systems with the emphasis on analytical and computer models. During his tenure at Kodak he spent three years in Japan and headed up the Imaging Processing Laboratory and Algorithm Developing Laboratory. He joined the University of Rochester in 1993 where he was the executive director of the Center for Electronic Imaging Systems and taught through the Computer and Electrical Engineering Department. He joined Sharp Laboratories of America in 2000 where he headed the Color Imaging Group. Dr Kriss retired in 2004 but is still active as a consultant, Adjunct Professor at Portland State University, IS&T activities, and as the Editor in Chief of the Wiley-IS&T Series on Imaging Science and Technology. Dr. Kriss was recently elected president of the International Congress of Imaging Science.

The Influence of New Light Sources on the Appearance of Paper and Print

Ole Norberg
Voxvil AB
Örnsköldsvik, Sweden
ole.norbeg@voxvil.se

Abstract

According to EU regulations, incandescent light bulbs and other energy inefficient lamps started to be gradually replaced by more energy efficient lamps in Europe. Today, lighting corresponds to almost 20% of all electricity consumption. By switching to more energy efficient lighting products, energy savings and reduced negative contribution to climate change can be achieved at the same time. When switching to energy efficient lamps there are two main alternatives; compact fluorescent bulbs or Light Emitting Diods (LED). As lighting changes so will the way we perceive objects and colour. Incandescent light bulbs, compact fluorescent bulbs and LED lighting all have different characteristics that will influence the way we for example recognize an object, experience a room or judge a colour. Furthermore, the whiteness level of a printing paper is considered as an important quality measure. High paper whiteness improves the contrast to printed areas providing a more distinct appearance of printed text and colors and increases the number of reproducible colors. As most paper and some printing ink include fluorescent material the present of UV-light is critical to the activation of these components. LED light sources have been showed to include only small amounts of UV-light thus the activation of fluorescent material be very low. The objective of this study was to evaluate how different lighting influence the way we perceive paper whiteness and also how it correspond to the calculated Whiteness level of paper.The results show that the appearance of the samples included in this study differs widely depending on the viewing situation. Some paper samples gain and some loose in appeared whiteness when the viewing conditions changes. One reason is the level of activation of the fluorescent whitening agent (FWA) in the paper. However, the activation of FWA does not explain all the difference. Other aspects on the appearance of whiteness and colour under different light sources must be identified and quantified.

1 Background

According to EU regulations, incandescent light bulbs and other energy inefficient lamps started to be gradually replaced by more energy efficient lamps in Europe. It is believed that similar regulations soon will be introduced in US. Today, lighting corresponds to almost 20% of all electricity consumption. By switching to more energy efficient lighting products, energy savings and reduced negative contribution to climate change can be achieved at the same time. The initial higher cost of energy efficient light bulbs is quickly regained, as these new lamps use significantly lower amount of electricity and last much longer. When switching to energy efficient lamps there are two main alternatives; compact fluorescent bulbs or Light Emitting Diods (LED). Compact fluorescent bulbs have been around for several years while LED-lighting has just entered the market. LEDs offers further advantages than compact fluorescent bulbs by being even more energy efficient, lasting longer and, maybe most important, they don't contain mercury. The expectation is that by 2020 nearly half of the commercial lighting business will be LED-lighting.

As lighting changes so will the way we perceive objects and colour. Incandescent light bulbs, compact fluorescent bulbs and LED lighting all have different characteristics that will influence

the way we for example recognize an object, experience a room or judge a colour. Furthermore, the whiteness level of a printing paper is considered as an important quality measure. High paper whiteness improves the contrast to printed areas providing a more distinct appearance of printed text and colors and increases the number of reproducible colors. As most paper and some printing ink include fluorescent material the present of UV-light is critical to the activation of these components. LED light sources have been showed to include only small amounts of UV-light thus the activation of fluorescent material be very low. The objective of this study was to evaluate how different lighting influence the way we perceive paper whiteness and also how it correspond to the calculated Whiteness level of paper.

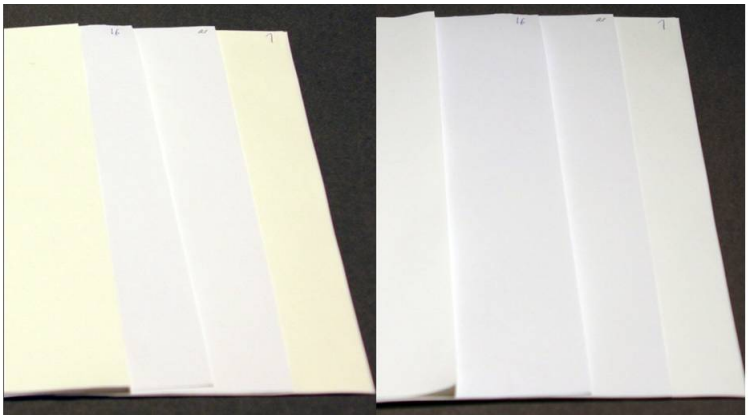


Figure 1: The effect of different light sources on the paper whiteness, compact fluorescent lamp (left) and LED lamps (right).

2 Method

In order to investigate the influence of different light sources on the whiteness appearance of paper a set of commercial papers with different whiteness levels were selected. The paper samples were evaluated by spectral measurements as well as by visual assessment studies. The whiteness values for the papers samples were calculated according to the CIE Whiteness specifications. Moreover, the samples were also characterized for respective viewing situation by spectroradiometer measurements.

This study include four different viewing situations, three lab studies with different light sources and one study representing an average office environment with a combination of light coming from fluorescent tubes and daylight through windows. For the lab studies, all light sources were chosen to produce the same colour temperature (3000K) and the same Luminance level (about 600 cd/m^2). The white LED light source used in this study is phosphor-based with a mix of GaN (peak at about 465 nm) and more broadband Ce3+:YAG phosphor (emits at roughly 500-700 nm).

Visual assessment study

The samples were presented as opaque pads (6 sheets) in A5 format. The samples were randomly coded from “A” to “P” with the code written on the back side of the pad. A panel of observers with mix in sex, age and of experience of judging paper was selected. Each observer was asked

Figure 2: The light sources used in this study, upper left: compact fluorescent, LED and Halogen.

to rank the samples according to whiteness, from lowest white to highest white. After ranking the samples, the pads were turned upside-down and the observer response could be recorded. For each observer, the sample ranked as lowest white has got the value 1 and second lowest value 2 and so on to the whitest sample getting the value 16. The average ranking was calculated by taking the mean value for each sample.

3 Results

Measurements of light sources

All light sources were characterized by spectroradiometer measurements 3. The expected spike of the fluorescent light sources is noticeable. Furthermore, it can be noticed that the amount of FWA-activating light, below 420nm, is very low for LED light source. The light source spectra were used as white point in the calculation of colour coordinates for the samples in respectively viewing situation.

Visual assessment

The panel of observers was a mix of "expert observers" and "non expert observers". The results for "expert observers" were compared to "non expert observers" and no significant different could be noticed. The ranking of the samples for different light sources are presented in 4. The variations in ranking between observers were in general low. The ranking for Fluorescent and Halogen light sources are similar with exceptions for a few samples that were proven hard to judge. For the LED light source the ranking is different and particularly some samples gain or lose in perceived whiteness.

The color coordinates of each sample were calculated for each viewing situation 5. The appearances of the samples differ widely between the different viewing situations. Under LED and Halogen light source the samples appear neutral or slightly reddish while under Fluorescent light source the samples appear bluish. It can also be noticed that the samples are wider spread under Fluorescent light.

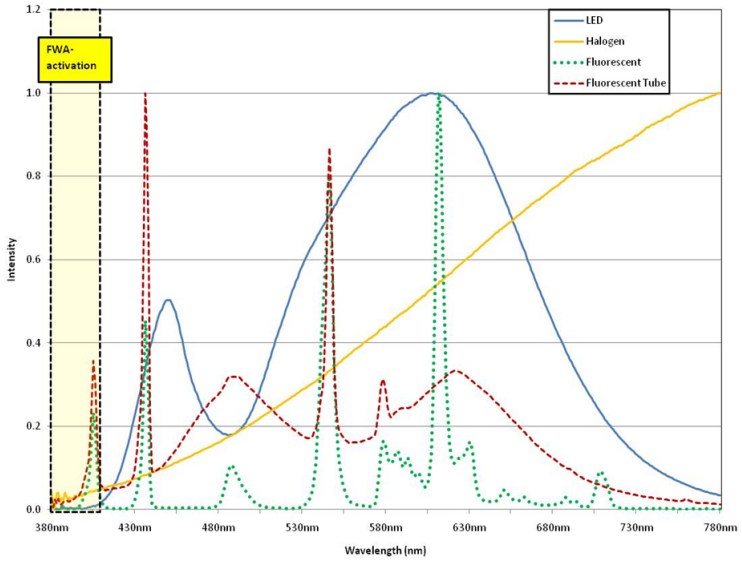


Figure 3: Spectral distribution of the light sources used in this study. The FWA is activated by light of wavelength shorter than 420nm. A Fluorescent tube is included in the graph as reference.

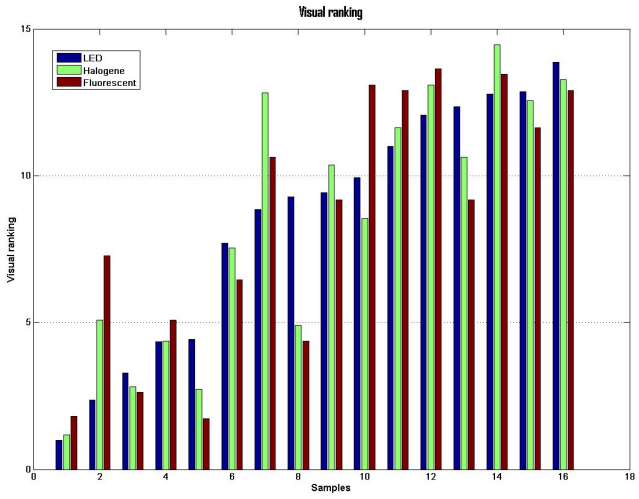


Figure 4: Visual ranking of samples under different light sources.

4 Conclusion

The appearance of the samples included in this study differs widely depending on the viewing situation. Many of the judgers in the visual assessment study comment on the large difference in appearance between the different viewing conditions. Also the split up of the samples differ between the light sources. The variation between the observers was smallest for the study under Fluorescent light. It can be explained by the split up of the samples in Figure 5, there are larger

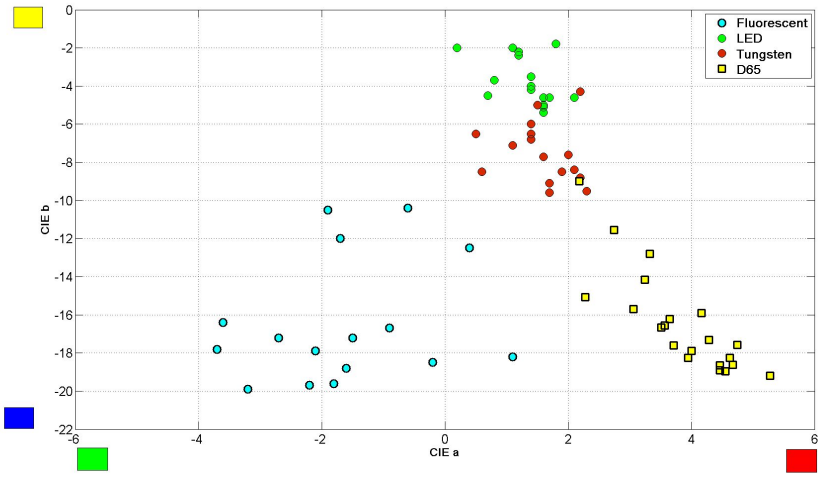


Figure 5: Calculated CIE $L^*a^*b^*$ coordinates for respectively viewing situation.

differences between the samples under that light condition.

Some paper samples gain and some loose in appeared whiteness when the viewing conditions changes. One reason is the level of activation of the fluorescent whitening agent (FWA) in the paper. However, the activation of FWA does not explain all the difference. Other aspects on the appearance of whiteness and colour under different light sources must be identified and quantified.

References

[1] European Comission Energy saving light bulbs, `http://ec.europa.eu/energy/lumen/index_en.htm`

[2] Light-emitting diode, `http://www.wikipedia.org`

[3] Pike Research, "Energy Efficient Lighting for Commercial Markets" (2010), `http://www.pikeresearch.com`

[4] Engeldrum. P. G., "Psychometric Scaling", Imcotec Press, Winchester, MA, ISBN 0-9678706-0-7, 2000

[5] Pauler N., "Paper Optics", AB Lorentsen & Wettre, 1998

[6] Hunt R.W.G., "Measuring Colour", Halsted Press, ISBN 0-470-20986-0, 1989.

[7] Norberg O., "Paper whiteness and its effect on the reproduction of colors", Proceedings of Human Vision and Electronic Imaging XII – IS&T/SPIE 19th Symposium on Electronic Imaging, San Jose, CA, US (2007).

Image Workflow and Applied Color Management in Police Investigations

Tino Günther
Landeskriminalamt Niedersachsen
Kriminaltechnisches Institut
Dezernat 55 - Bildtechnik
Schützenstr. 25
30161 Hannover
tino.guenther@polizei.niedersachsen.de

Abstract

The presentation "Image workflow and applied color management in police investigations" explains how images are processed at the police in Lower Saxony in Germany.

If a policeman takes a photo anywhere in Lower Saxony, it will be sent with the help of special computers called "BASt" to the State Criminal Investigation Department of Lower Saxony. Nearly every police office is equipped with one of the 160 BASts. With this machine, the policeman is able to choose which pictures and how many of them he needs as a print-out.

The incoming orders are automatically inserted into the database with the aid of the digital workflow manager. The next step includes the issue of an order sheet which is needed for the following image processing and printing. With a program called "filter client", color correction is made by hand to get more information from the images.

The print-outs are produced on an AGFA chemical lab or a Noritsu dry lab, finally checked out and sent to the recipient. After being checked out, the images are available at the scene management system.

At the scene management system the policeman is able to order an image report, which is also produced at the State Criminal Investigation Department within a color management environment. The program "Canoflow" is used for managing the image report orders.

Color management is an essential aspect in image processing and printing in police investigations. Many different machines and technologies are used, for example monitors with low and wide gamut, inkjet, xerography, thermo-sublimation and chemical minilabs. It's expected that every machine reproduces the same color.

We researched for a digital printing standard. We couldn't find a paper for this, so we preferred orientation to "ISO 12647-7". This has the advantage of some corresponding standards like "ISO 3664", the given color spaces "ISO coated v2", "eciRGB v2" and the measuring conditions.

Recent Work on Fluorescence in Colour Management

Philip John Green
Adjunct Professor
Gjøvik University College
philip.green@hig.no

Abstract

A number of recent projects on fluorescence in colour management were summarised.

Substrates with high levels of fluorescent whitening agents are widely used in graphic arts to increase the perceived brightness and whiteness of the substrate. FWA molecules are excited by photons at UV wavelengths and emit in the visible region of the spectrum. The efficacy of this fluorescence is less stable than the ink and paper reflectance, leading to a change in colour over time especially when the material is exposed to high levels of illuminance of UV content. Production papers usually have high FWA levels, while proofing papers typically have much lower levels.

The core of the problem in colour reproduction is that the amount of excitation (and hence the spectral radiation of the illuminated material) depends on the UV content of the illumination, and this varies significantly in both viewing conditions and the light sources used in measurement instruments. Since different substrates have different levels of FWA, this makes it a difficult problem to match the appearance of a white point, and to reproducibly measure it.

A spectral match between different substrates cannot be achieved, and when a colorimetric match is targeted by minimizing the absolute colorimetric difference between a production paper and proof, the proof tends to be printed with a very bluish white point. As the match is not spectral, it will only hold under a given illumination source spectral power.

Three approaches can be used to address these problems.

The first approach is to suppress fluorescence at the measurement stage, so that there is no UV-excited blue emission in the measurement data. This is a simple solution works reasonably well if observers are fully adapted to the white point of each medium. However, it prevents an exact match between different media.

The second approach is to target an exact colorimetric match with fluorescence included. This approach is supported by ISO 13655:2009 M1 condition (which requires that the instrument source matches the D50 spectral power distribution, including the UV portion from 310nm) and ISO 3664:2009 which also requires a spectral power of D50 in the illumination source.

The third approach is to achieve a match between different media through soft proofing. In this case, if the colour appearance of the target medium can be predicted under different amounts of UV the soft proof can accurately show its appearance.

Recent work at London College of Communication which addressed these different approaches include:

- Loeffler and Green, 2008 [6]. 'Scalable Model Spectra Enabling Estimation of Fluorescence Characteristics of Brightened Offset Papers Under Different Illuminants' made bi-spectral measurements of a sample of OBA-containing papers, derived fluorescent identifier for papers, modelled absorption and emission spectra, and modelled fluorescent emission.

A Pragmatic Approach to Colour Input Device Characterisation

Lindsay W. MacDonald
Photogrammetry, 3D Imaging and Metrology Research Group
Department of Civil, Environmental and Geomatic Engineering
University College London
London WC1E 6BT
lindsay.macdonald@ucl.ac.uk

Abstract

It has been well known for at least 35 years that colour space transformations from device signals (such as R,G,B) to device- independent values (such as CIE L*a*b*) can be effectively implemented by a lookup table (LUT). To overcome limited memory the most significant bits (MSB) address the nearest neighbours in a smaller table, and the least significant bits (LSB) interpolate between them. If the functional relationship between input and output is known then the cells in the table can be loaded with exact values. More commonly some sort of functions (typically third-order polynomials) are fitted by regression to a limited number (typically ¡ 200) of measured samples, and the polynomial is interpolated to load the LUT.

A new technique has been developed for cases where the spectral sensitivity of the device is accurately known a priori (e.g. response at 10 nm wavelength intervals). The input and output values are calculated for each a huge number (¿100,000) of synthetic but realistic reflectance functions, and the results used to populate all the cells in the LUT. Statistical techniques are then used to fit and smooth the transfer function, giving much higher accuracy than simple polynomial functions.

Spectral Goniophotometry

Project Description and Initial Results

Niklas Johansson
Department of Information Technology and Media, Digital Printing Center
Mid Sweden University
Örnsköldsvik, Sweden
niklas.johansson@miun.se

Abstract

A spectral goniophotometer will be used to examine topics such as anisotropy of reference materials, separation of surface scattering from bulk scattering and the influence of measurement geometry on reflectance measurements. Initial results show that the absolute measurement capability of the instrument is very good, and approximations of measurement geometries with diffuse illumination by numerical integration of a set of bidirectional measurements show promising results.

1 Introduction

A goniophotometer is an instrument that can measure the amount of light reflected or transmitted in a certain direction for different angles of incidence, and spectral goniophotometers can carry out these measurements spectrally resolved. Angularly resolved reflectance and transmittance measurements are of great interest in many scientific fields. Goniophotometric measurements are a necessity when characterizing so called gonio-apparent materials, i.e. materials that exhibit a shift in color depending on the illumination and viewing angle. Coatings with pearlescent pigments, such as metallic paints used within the automotive and packaging industry are typical examples of gonio-apparent materials. Goniophotometric measurements are also heavily used within computer graphics, for realistic rendering of illuminated surfaces.

Light scattering in turbid media, such as paper, is one topic where the spatial distribution of the reflected light is of particular interest. In order to strengthen the knowledge in this field and to further develop the numerical simulation tools that are currently used within our research group, a spectral goniophotometer has been acquired at the Digital Printing Center in Örnsköldsvik. With this instrument, many interesting quantities can be measured and in particular the bidirectional reflectance distribution function (BRDF) [4]. The BRDF is a fundamental quantity from which the reflectance under any geometric configuration can be derived by suitable integration. The reflectance of a material is fully described by the BRDF and knowing the BRDF of a material allows us to describe the visual appearance of it under any illumination in terms of spatial and spectral distribution of the reflected light. However, for a complete description of the appearance, physiological and psychological effects must be taken into account.

2 Instrumentation

The instrument consists of a Perkin-Elmer Lambda 1050 UV/VIS/NIR spectrophotometer with ARTA goniophotometer accessory from OMT Solutions BV, see figure 1. Some of its key features are:

- Wavelength range 220 -2500 nm. Double monochromator for high spectral resolution of 0.05 nm in UV/VIS and 0.20 nm in NIR.
- Double beam system for absolute measurements, where the reference beam sent directly to the detector via optical fibers.
- Tilt stage for out-of-plane measurements at azimuthal angles ± 75 deg.
- Equipped with a polarizer in the sample beam for s- and p-polarized incident light.
- Large angular range, 10deg -350 deg, enabling both reflectance and transmittance measurements, figure 2.
- Variable detector aperture for adjustable detection solid angle in the range 0.06 - 0.003 sr^{-1}

(a) Goniophotometer (b) Beam path

Figure 1: ARTA accessory installed in Perkin Elmer spectrophotometer 1a. The beam path in a reflectance measurement is shown in 1b. The sample (2) is illuminated by monochromatic light from the light periscope (1) and reflected into the integrating sphere detector (3). Reprinted by permission of OMT Solutions BV.

3 Planned activities

First of all, the absolute measurement capability of the instrument, utilizing the double beam configuration, will be examined. Measurements of reference standards, with reflectance values traceable to Metrology institutes such as the Institute for National Measurement Standards at the National Research Council of Canada (INMS-NRC) will be carried out. The instrument will also be compared with existing reference instruments in an inter-instrument comparison study, where different paper grades are measured at the participating institutions.

In the paper industry, d/0 is the de facto standard measurement geometry for reflectance measurements, using diffuse illumination and measuring along the normal of the sample at zero degrees. Measurement geometries with diffuse incident or reflected light cannot be carried out

(a) Transmittance setup (b) Reflectance setup

Figure 2: Top view of the ARTA sample compartment. Instrument setup for transmittance measurement 2a and reflectance measurement 2b. Reprinted by permission of OMT Solutions BV.

directly with a spectral goniophotometer. Instead, a set of bidirectional measurements need to be carried out, and subsequent integration gives an estimate of the d/0 or 0/d reflectance [2, 3]. The degree of accuracy is unknown and will be studied in this activity. Further on, the capability of measuring gloss according to standardized methods will also be tested.

Characterization of different reference materials is another topic that will be investigated, emphasizing on anisotropic scattering (deviation from Lambertian behavior). The general assumption that reference materials are Lambertian, and using only one bidirectional reference value when normalizing relative measurements carried out at other geometrical configurations, introduce errors in measured reflectance values. The magnitude of this error is unknown and will be examined further. Different materials applicability as reference materials for different paper grades, but also for other types of materials such as carpet, textiles, fur/hair, ceramics and plastics, is another question that will be addressed.

We will also look into the problem of how to separate surface reflection from bulk reflections. Measurements on different paper grades will be valuable for improving the advanced simulation methods developed within out research group. Models for simulating the bulk reflectance are already in use [1], and the next step is to include surface models. In this activity, we will examine existing surface models and their applicability, with focus on using different surface roughness measures as input parameters to the models. Later on, methods like ellipsometry will also be studied.

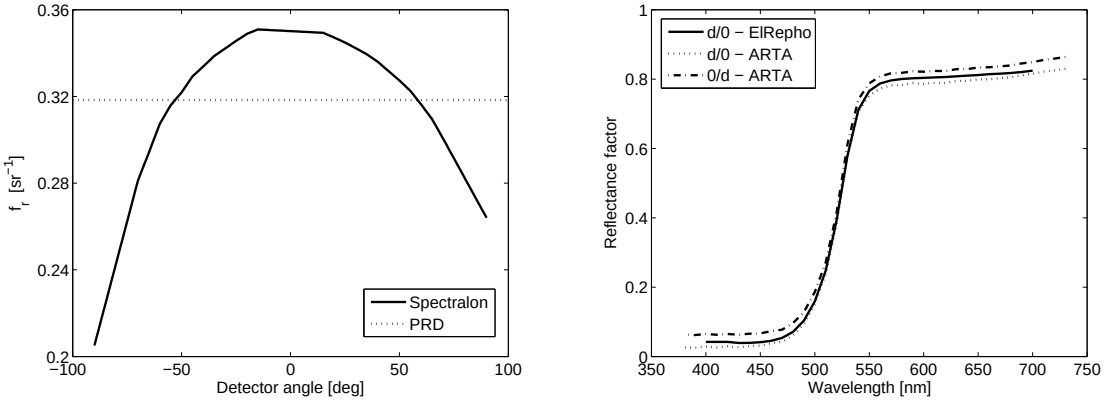
An anisotropy study of fluorescence will also be undertaken, using a set of experimental papers with varying basis weight, amount of filler and amount of fluorescent whitening agent. Angularly and spectrally resolved reflectance measurements will be carried out with the goniophotometer. The measurements at some bidirectional geometries will be complemented with bispectral measurements as well.

Finally, guidelines for how to design an instrument adapted for industrial use, with main emphasis on reducing the amount of bidirectional measurements needed to characterize a sample, will also be developed. Measuring the full BRDF with a scanning instrument, where the measurements at different angles are taken one by one in a sequence, is very time-consuming. For practical use in industrial laboratories, reducing the measurement time is a necessity. As

an alternative to scanning systems, camera-based instruments offer faster measurement times since they can capture several measurement angles in a single exposure. They are, however, in general less accurate than scanning systems.

4 Initial results

So far, most of the work has been focused on calibration issues and evaluation of the instrument regarding precision and absolute measurement capability. Figure 3a shows absolute measurements of a Spectralon diffuse reflectance standard, with an 8/d-reflectance value traceable to National Institute of Standards and Technology (NIST). Spectralon, which is made of pressed PTFE and a trademark of LabSphere, is known for its high reflectance and excellent Lambertian behavior. The measured 8/d-reflectance factor, by numerical integration of the bidirectional measurements and assuming azimuthal symmetry, is 0.998 and the tabulated reference value is 0.991, which gives us a relative error of 0.7%.



(a) Absolute measurement of reference material

(b) d/0 and 0/d measurements with ARTA

Figure 3: Absolute measurement at 520 nm of Spectralon diffuse reference standard. The illumination angle is 8 deg and the polar detector angle is varied from −75 deg to +75 deg in steps of 5 deg. Values outside this interval have been extrapolated 3a. Absolute d/0- and 0/d-measurements of a yellow NCS patch (NCS S0580-Y), together with corresponding measurement with a commercial d/0-instrument (Elrepho), is shown in 3b.

In figure 3b, both a d/0- and a 0/d-measurement is shown and compared against a measurement with the Elrepho instrument, which is the standard d/0-instrument used in the paper industry. The color difference, expressed in CIE 1976 ΔE^*_{ab} metric [5] , between ARTA d/0 measurement and Elrepho is 3.6 ΔE^*_{ab}, while the difference between ARTA d/0 and ARTA 0/d is 11.6 ΔE^*_{ab}. It should be noted that the goniophotometer measurements has not been adapted to try to take the gloss trap of the Elrepho instrument into account.

References

[1] Per Edström. A Fast and Stable Solution Method for the Radiative Transfer Problem. *SIAM Review*, 47(3):447–468, 2005.

[2] Xiaofan Feng, John R. Schott, and Timothy Gallagher. Comparison of methods for generation of absolute reflectance-factor values for bidirectional reflectance-distribution function studies. *Applied Optics*, 32(7):1234–1242, 1993.

[3] Silja Holopainen, Farshid Manoocheri, Erkki Ikonen, Kai-Olaf Hauer, and Andreas Höpe. Comparison measurements of 0:45 radiance factor and goniometrically determined diffuse reflectance. *Applied Optics*, 48:2946–2956, 2009.

[4] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis. *Geometrical Considerations and Nomenclature for Reflectance*. Number NBS monograph 160. National Bureau of Standards, 1977.

[5] Gaurav Sharma. *Digital Color Imaging Handbook*. CRC Press, 2003.

Optimal Wavelengths of Colour Laser Scanners

Lindsay W. MacDonald
Photogrammetry, 3D Imaging and Metrology Research Group
Department of Civil, Environmental and Geomatic Engineering
University College London
London WC1E 6BT
lindsay.macdonald@ucl.ac.uk

Abstract

A colour laser scanner represents the extreme case of colour acquisition, insofar as it samples the surface reflectance of an object at three single wavelengths. It is therefore has the greatest possible degree of metamerism of any colour image capture system. Suppose that any wavelength could be chosen for each of the long, medium and short wavelength (R,G,B) lasers? What wavelengths would be optimum in the sense of minimising the colour errors for a large set of typical samples? The surprising answer is that the optimum wavelengths are almost invariant of samples or colour difference metrics, in the ranges 450-460, 530-535 and 590-600 nm. These are close to Thornton's 'prime wavelengths of vision' and indicate something fundamental about the human visual system.[1]

[1]The full paper has been published in Nineteeth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications, San Jose, California; November 2011; p. 357-; ISBN / ISSN: 978-0-89208-297-1

Spectral Model of an Electro-Photographic Printing System

Michael A. Kriss
MAK Consultants
506 NE 193RD Avenue
Camas, WA 98607, USA
makriss@comcast.net

Abstract

At EI 2007 in San Jose, California detailed physical models for monochrome and color electro-photographic printers were presented. These models were based on computer simulations of toner-dot formation for a variety of halftone structures. The optical interactions between the toner-dots and the paper substrate were incorporated by means of an optical scattering function, which allowed for the calculation of optical dot-gain (and physical dot-gain) as function of the halftone structure. The color model used simple red-green-blue channels to measure the effect of the absorption and scattering properties of the cyan, magenta, yellow and black toners on the final half-tone image. The new spectral model uses the full absorption and scattering spectrum of the image toners in calculating the final color image in terms of CIE XYZ values for well-defined color and gray patches. The new spectral model will be used to show the impact of halftone structure and toner-layer-order on conventional dot-on-dot, rotated dot and error diffusion color halftone systems and how to minimize the impact of image toner scattering. The model has been expanded to use the Neugebauer equations to approximate the amount of cyan, magenta, and yellow toners required to give a "good" neutral in the rotated dot halftone and fine tuning is achieved by adjusting the development threshold level for each layer to hold a good neutral over the full tonal range. In addition to the above fine-tuning, cyan, yellow and magenta offsets are used to find an optimum use of the halftone dither patterns. Once a "good" neutral is obtained the impact on dot gain, color reproduction and optimum layer order can studied with an emphasis on how the full spectral model differs from the simpler three-channel model. The model is used to explore the different approaches required in dot-on-dot, rotated dot and error diffusion halftones to achieve good results.

1 General Results

The spectral model for a color electro-photographic printer has been developed that allows one to study the nature of different halftone screens under the conditions of differing scattering and absorption spectra. The model ranges from 380 to 730 nanometers and allows one to make color reproduction calculations in terms of the CIE XYZ values and it various transforms. The results of the spectral model agree with the RGB channel model in terms of determining the optimal layer order for a specific set of toner scattering spectra in that the worst and best layer orders are the same, but the spectral model shows greater sensitivity to changes in layer order than the three-channel model. The biggest difference between the two models is that the net dot-gain predicted by the three-channel model (using unwanted absorptions) is much higher than that predicted by full spectral model; the three-channel model predicts a Yule-Nelson value of n=1.7 while the spectral model predicts n=1.4. In the rotated dot halftone screen the solving of the Neugebauer equations is very useful in defining the initial level for the cyan, magenta and yellow toners to maintain a good neutral, while in the dot-on-dot halftone screen it is more important to keep the cyan, magenta and yellow dots the same size. In all cases, adjusting the exposure threshold in the individual layers is key to maintaining a good neutral

scale as are slight offsets in the cyan, magenta and yellow values optimized for the given halftone dithers. The spectral model shows the detailed shifts in color reproductions for a color chart as a function of the layer order and scattering spectra, thus allowing for optimum system design.

2 Computer model background and new spectral properties

The physical models presented at EI 2007 [3][4] are the basis for the current spectral model presented below. The earlier monochrome model focused on the formation of the electro-photographic pigment dot, the halftones structures, the physical dot-grain defined by the physical growth of the halftone dot due to exposure and development, and the optical dot-grain defined by the coupling of the optical spread function of the scattering paper support. The model agreed with trends of other theoretical and experimental results [5][1]. The color model assumed a simple four layer systems consisting of cyan, magenta, yellow and black halftones defined by their peak and unwanted absorptions and their scattering properties defined by the percent scattering in well defined red, green and blue bands. The absorptions for each layer were characterized by their peak absorptions, red for cyan pigments, green for magenta pigments and blue for yellow pigments, and the unwanted absorptions such as the blue absorption of the cyan pigment, etc. The scattering model traced the light scattered by each pigment back toward the air-pigment layer interface, the light transmitted and absorbed into the layered structure, light reflected by the paper support, and took into account the multiple reflections of the scattered and transmitted light back into the pigment layers. The net result was a measure of the red, green and blue light reflected by the color halftone. This model also provided a measure of the optical and physical dot-gains as well as the color shifts introduced by the scattered light as a function of the layer order. For example, if the yellow pigment scattered red light, the impact of the scattering was greater when the yellow layer was on top versus when it was below the cyan layer where the scattered red light was absorbed before it reached the pigment-air interface and into the net, measured reflected light.

The current spectral model, developed in Mathematica®, replaces the red, green and blue channels with the full spectral range, 380 nanometers to 730 nanometers, resulting in a full spectrum of the net reflected light, which in turn could be studied by either its XYZ values using the CIE Color matching functions or simple red, green, blue filter sets (like Status A filters used to look at prints and transparencies). Each halftone layer essentially has 36 parallel layers, one for each wavelength of light in the range defined above. As in the previous models, each pigment dot is made up of a 21×21 array and each color or neutral patch is made up of a 21 x 21 array of dots (a 441×441 array of pixels). This ensures that an accurate measure is calculated for the final spectral reflection of any given patch. In addition to the scattering and absorption properties of the various pigments, spectral front surface reflection curves are defined for clear areas (paper support) and pigmented areas. Also, a viewing illuminant can be defined to observe and calculate the shift in color reproduction (not assuming any color adaptation of the human visual system.) In order to make "easy" comparison, it is important to start with a "good neutral" for the case of no scattering in an "ideal" hard dot halftone. To achieve this models take into consideration the unwanted adsorptions and adjusts the cyan, magenta and yellow coverage to obtain a neutral scale. An alternate method to get the appropriate balance between the cyan, magenta and yellow pigments is to solve the spectral Neugebauer Equations for the required cyan, magenta and yellow coverage for a neutral. Deviations form the neutral due to dot-gain and scattering are easy to detect.

3 Selected comparisons with the non-spectral color model

The added value of the spectral model is found by comparing its results to the non-spectral models. To do this the following steps were taken. The first step is to define the four pigments, cyan, magenta, yellow and black, to be used for both the spectral and non-spectral models. Each pigment will have an absorption and scattering spectrum. Figure 1 shows the absorption spectra for four such pigments. Note that the cyan, magenta and yellow pigments are selected such that a uniform coating of all three would produce a unit neutral (of reflected light) of density 2 or 0.01 reflectance. Figure2 Shows the percent scattering from each pigment. Note that this is the amount of light scattered upwards from the pigments. Any light scattered latterly or down toward the paper support is considered to be transmitted light, which may or may not be absorbed by the layers below. Since the non-spectral model uses red, green, blue bands of absorption and scattering, each of the spectral curves for absorption and scattering are "measured" with three pass bands of radiation as follows: 390 to 490 nanometers for blue, 500 to 600 nanometers for green and 610 to 710 nanometers for red. Applying these filters to each of the spectral curves provides the averaged values of peak and off-peak absorption and scattering required by the non-spectral model and ensures that the same pigments are used for both models.

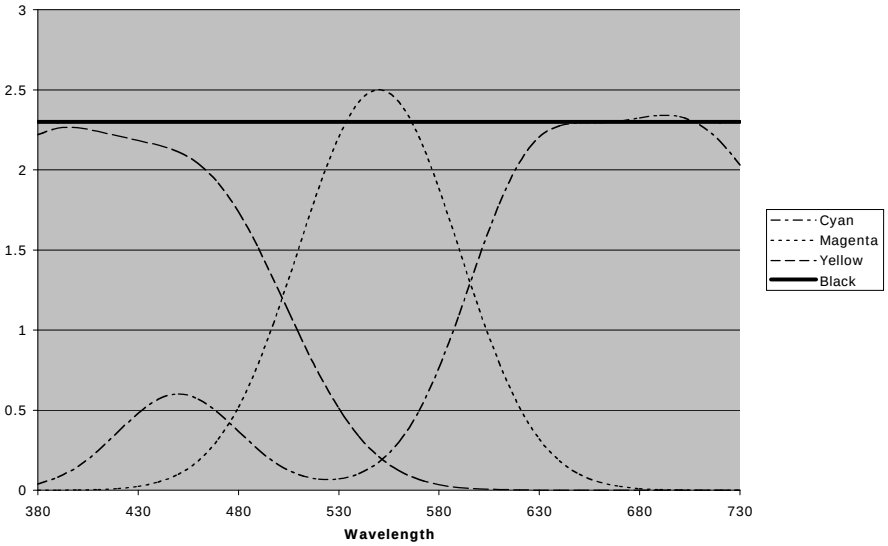


Figure 1: The four tone absorption spectra.

The first comparison uses a simplified case of uniform coatings of pigments (no halftone structure) to establish the optimum layer order for the cyan, magenta, yellow and black layers given their scattering and absorption properties. The spectral and non-spectral models are identical in all facets except the non-spectral model uses the band pass filter obtained absorption and scattering averages, while the spectral model uses the full ten nanometer resolution data. The non-spectral model results in red, green and blue values for an input of a gray level of 128. The resulting red, green and blue values are plotted on a tri-chromatic diagram where the red, green, and blue axes are 120 degrees apart. The measure of "goodness" is the amount of chromatic shift from the neutral for each of the 24 combinations of layer orders. In the

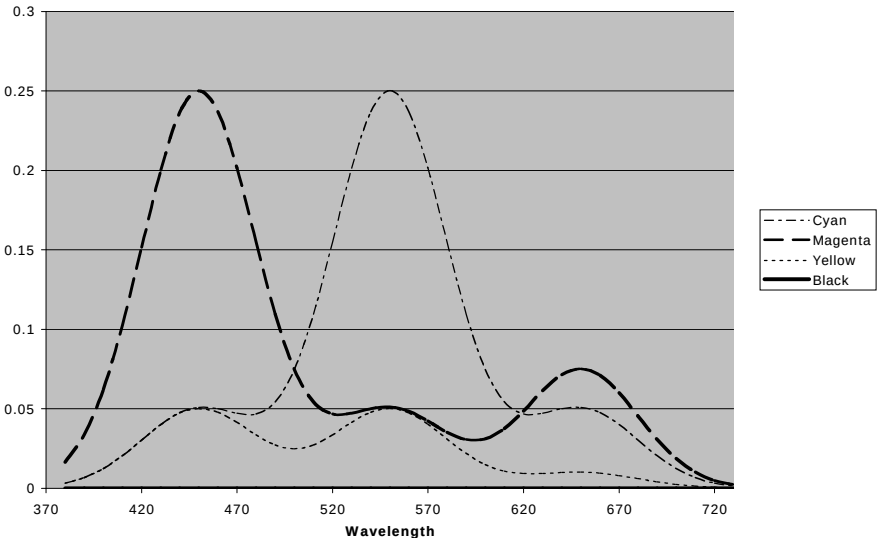


Figure 2: The light scattering spectra of the four-toner pigments.

case of the spectral model the X, Y, and Z values are calculated (assuming an equal energy illuminant) using the CIE Color Matching Functions and from them the chromatic deviation from a neutral are calculated for the same 24 combinations of layer orders. In each case the smallest deviation from a neutral is defined as the best layer order and the largest deviation form neutral is defined as the worst layer order. In the first example it is assumed that the cyan layer has strong scattering in the green region of the spectrum. The non-spectral model predicts that the best layer order is {black, magenta, cyan, yellow} where black is on top. The spectral model predicts that the best layer order is {black, yellow, magenta, cyan}. Note that in both cases the green light scattering cyan layer is below the magenta layer where the scattered green light will be absorbed. The non-spectral model show a 21:1 ratio between the best and worst chromatic deviation, while the spectral model show a 32:1 deviation. An analysis of the results from both models indicate that as long as the cyan layer is below the black and magenta layers, the chromatic shift due to the scattering of green light is small, thus allowing for a range of acceptable layer orders. Based on this one example, both models give similar results, but the spectral model gives a finer differentiation between each layer order.

A more challenging example is when the cyan pigment scatters substantial green light and the magenta pigment scatters an equal amount of blue light. Both models give the best layer order as {yellow, magenta, cyan, black}. The analysis of the data shows that for both models the cyan layer must be below the magenta layer and the magenta layer must be below the yellow layer for a very low chromatic shift from neutral. Likewise, both models show the worst layer order to be {cyan, yellow, black, magenta} where the scattering from the cyan layer is not attenuated. The layer order {cyan, magenta, black, yellow} has a chromatic shift close to that of {cyan, yellow, black, magenta} and represents a case where the scattering of the cyan and magenta layers are not very attenuated by the layers above them. Again, the spectral model shows a much greater range of chromatic shifts and sensitivity to layer order.

The results for these simple uniform layers of pigments indicates that both models give equivalent information and since the computational time of the spectral model is equivalent to

that on the non-spectral model, the spectral model should be used to determine the best layer order.

The next set of comparisons deal with full halftone models. Two halftones will be considered; the rotated dot where the black halftone is at 45 degrees, the yellow halftone at 90 degrees, the cyan halftone at –15 degrees and the magenta halftone at +15 degrees, and a dot-on-dot halftone using the 45 degree structure. The results will cover both the spectral and non-spectral cases, with and without scattering. In what follows the term "ideal" will include results that assumes a perfectly hard halftone dot (the ideal case) with no dot-gain due to the electro-photographic process and the term "real" that includes the electro-photographic process. Both the "ideal" and "real" cases include all the absorption and scattering effects. Space does not allow a set of color images of the resulting halftones or tone scale curves. Instead a general description of the results will be given. One diagnostic measure will be the Yule-Nielson "n" value [1] used widely to measure the dot-gain of halftone systems. The models (both the spectral and non-spectral models) provide as output a detailed look at the halftone patterns for each tone scale patch (with reflection values ranging from 1 to 0 in 9 or 16 uniform steps), the overall tone scale measured with Status A red, green and blue filters (in the spectral case) and simple R-G-B values in the non-spectral case, and a color map in the CIE x-y space for the spectral model and in an equivalent r-g space for the non-spectral model. The Yule-Nielson "n" values is calculated from the "average" of the "real" red, green and blue tone scale curves relative to the average of "ideal" red, green and blue tone scale curves.

The examples that follow come from the "best" and "worst" layer order from the results shown above. The "best" layer order is {yellow, magenta, cyan, black} and the "worst" layer order is {cyan, yellow, black, magenta}. Due to the statistical nature of the development model, each run of the model results in a different noise patterns for the halftone generation, hence gives slightly different results. A series of six runs of the same model parameters was run to check the variation on the non-spectral model. The variance is best summarized in the Yule-Nielson "n" value that had a range from $n = 1.60$ to $n = 1.67$ with a mean value of $n = 1.64$ and a standard deviation of 0.03. Thus any changes within this standard deviation must be considered due to the noise and not due to any meaningful differences in the inputs to the model.

Layer Order	Spectral Model		Non-spectral Model	
	Scattering	No scattering	Scattering	No scattering
C Y B M	1.40	1.33	1.66	1.67
Y M C B	1.43	1.43	1.62	1.63

Table 1: Yule-Nielson "n" values for the best and worst layer orders.

Layer Order	Spectral Model		Non-spectral Model	
	Scattering	No scattering	Scattering	No scattering
C Y M B	Shift to green	Slightly warm	Shift to green	Slightly warm
Y M C B	Slightly green	Slightly warm	Slightly warm	Slightly warm

Table 2: Color shifts in gray scale for the worst and best layer orders.

Table 1 summarizes the variation in the Yule-Nielson "n" value for the rotated dot halftones. It is clear that the non-spectral model gives a higher "n" value than the spectral model. The variations within each model indicate that there is little or no dependence on the "n" value

on the amount of scattered light. Table 2 describes the general color shifts in the gray scales for each model. It is clear that in both the spectral and non-spectral models that the impact of layer order is crucial in lessening the impact of the scattered light. When scattering occurs there is a strong shift to the green for the {cyan, yellow, magenta, black} layer order while only a slight shift to the green in the optimum {yellow, magenta, cyan, black} layer order. When no scattering is allowed, both models show slightly warm (slight shift to the red) gray scales. This indicates a slight modification in the halftone generation is required to slightly increase the amount of cyan pigment generated in the electro-photographic development process.

The dot-on-dot halftone system shows similar results as given in Table 1 and Table 2 for the rotated dot halftone. The spectral model gave a nominal value of $n = 1.4$, while that for the non-spectral model gave a nominal value of $n = 1.6$. The color shifts were similar, again indicating that the layer order is important in the presence of scattered light. The dot-on-dot halftone is much more sensitive to variations in dot development in the four pigment layers. If, for example, the cyan layer develops slightly larger dots, the neutrals will become more cyan. This implies that one could help compensate for scattering by adjusting the development of some combination of the layers to regain a true neutral. However, this also implies that there must be very tight control on the toner pigments, for if they change their spectral nature then the development process will have to be changed to compensate for this change.

Similar experiments were run on a nine (and sixteen) color array (red, green, blue, cyan, yellow, magenta, orange, purple, gray) and the results were similar to those above. If the worst layer order is used the colors shift toward the green, while if the best layer order is used or if there is not scattering, the colors reproduce well (but not perfectly).

The spectral model was also applied to a conventional error diffusion halftone. The results demonstrate the large dot gain found with all stochastic halftones. In this case n=2.47. There was a strong shift to the green when the layer order went from the optimum {Y, M, C, B} to the worst order {C, Y, B, M} showing that the scattering effects are independent of halftone structure.

4 Balancing the Gray Scale

One major goal of any halftone system is to produce a good gray scale. In all the previous work above, the gray scales tended to be slightly warm or showed slight swings between warm and cold, from level to level, when the light scatter was suppressed by using the optimum layer order. There are several system parameters that can be used to optimize the gray scale. The first is to use offsets in the cyan, yellow and magenta layers to best align the steps provided by the three-color halftone dither patterns. The second is to adjust the threshold development levels to control the dot size for the cyan, yellow and magenta layers. The third is to make a better measure of the impact of the unwanted absorptions of the three-color pigments. In the earlier attempts to adjust the amount of cyan, yellow and magenta pigments, Status A type filters were used to measure the unwanted absorption for the three pigments. This results in a simple 3×3 matrix relating the amount of red, green and blue reflectance to the amount of cyan, yellow and magenta pigments. This matrix can be inverted to give the amount of cyan, magenta and yellow pigments to the amount of measured red, green and blue reflected light. Using equal amounts of red, green and blue outputs, one can calculate the amount of cyan, magenta and yellow pigment required. For the case of the pigment spectra used in this paper, the ratio of cyan to magenta to yellow pigment was found to be {1: 0.86: 0.9}. Thus the original amount of magenta pigment (255-green code input) was multiplied by 0.86 and a similar calculation was made for the yellow pigment. This is due to the higher unwanted absorptions of the cyan

pigment (thus less magenta and yellow pigments are required.) However, this resulted in warm gray scales when the scattering was "turned off" or suppressed by the proper layer order. To get a better measure of the impact of the unwanted absorptions the spectral Neugebauer Equations [2] were solved (using Mathematica®). The solution is to require equal red, green and blue values (Status A filter values) and to solve for the appropriate cyan, magenta and yellow coverage. The solution is non-linear and gives several answers, many of them complex and other with no physical meaning. However, there was always one where the three coverage values were always positive and less than one; this was the solution that was used. For the three pigments used in the study the ratio of cyan to magenta to yellow was {1: 0.7, 0.7}. This resulted in more cyan coverage to magenta and yellow coverage than found when using the unwanted absorptions as outlined above. The results, while better, still showed a slightly warm gray scale. Further improvement was sought by adjusting the level of cyan pigments by adding a positive offset to the cyan coverage and increasing the cyan dot size by lower the development threshold. The resulting tone scales are show in Figure 3 and Figure 4 below.

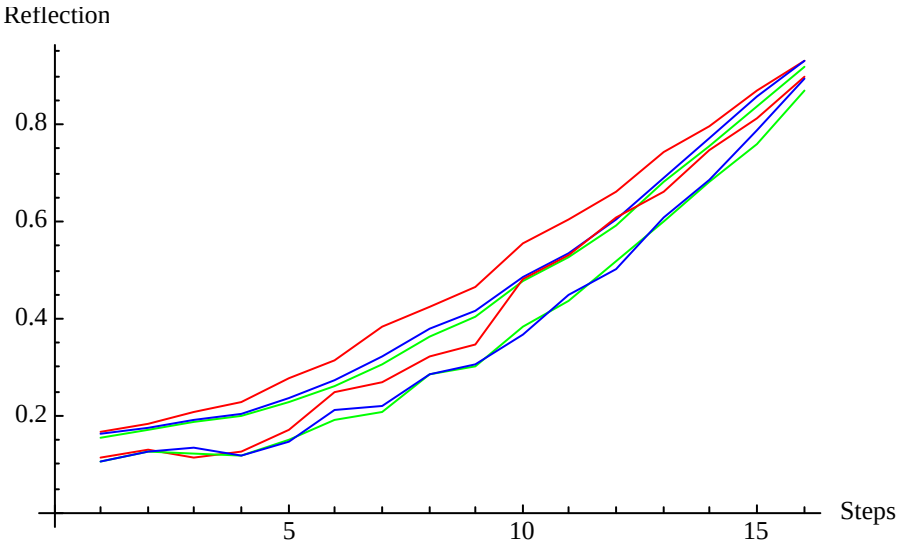


Figure 3: The nominal "ideal" and "real" tone scale. The top curve set is the "ideal" tone scale and the bottom curve set is the "real" tone scale.

Further manipulations of the three offsets and dot size can improve the gray scale, but larger dither patterns may be required to "smooth out" the overall tone scale.

5 Conclusions

The spectral model, while more complex and computationally intensive, produced more accurate differentiation in color shifts. The non-spectral models produced similar color shifts to those of the spectral model, but consistently gave higher Yule-Nielson "n" values for dot-gain. It is felt that the spectral model's values for "n" are more accurate, and thus the spectral model should be used when trying to understand the nature of dot-gain, while either model can be used when trying to predict color shift due to varying spectral scattering curves for the toner pigments. Future work will expand the spectral model to other halftones systems, developing

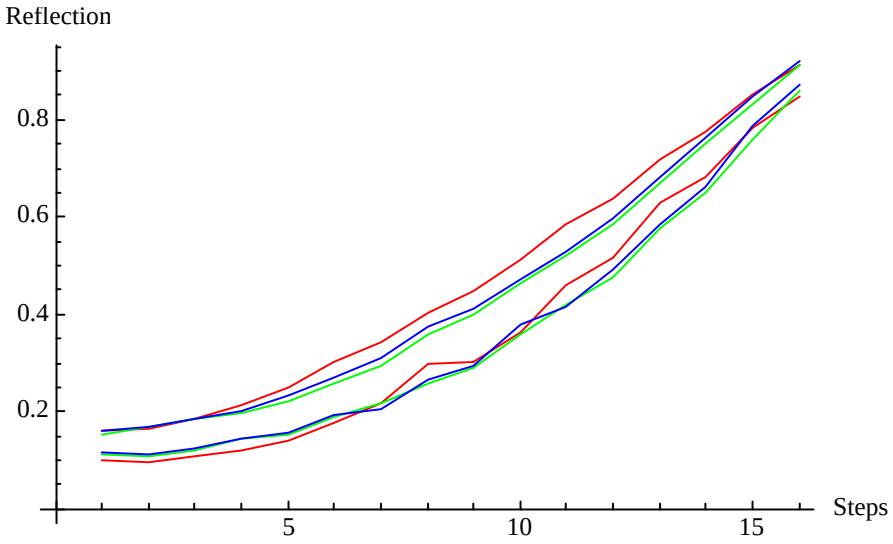


Figure 4: The "real", top, and "ideal", bottom, tone scales when a positive cyan offset is used and a larger cyan dot is appled to the halftones. Note that the curves are better mathched, but not perfect, and they have moved apart indicating that the larger cyan dot size has increased the amount of dot gain.

color tables for a well-defined system and applying the concept to "real" images.

The full spectral model provides a very good teaching tool to understand the often complex interactions in an electro-photographic printing system.

Acknowledgements

I wish to thank Dana Smith of the Sharp Laboratories of America for his help in formulating these models and his invaluable experimental measurements to confirm the accuracy of the models.

References

[1] S. Inoue, T. Tsumura, and Y. Miyake. A new model to estimate optical dot gain in printing and its applications. In R. Eschbach, editor, *Recent Progress in Digital Halftoning II*. IS&T, 1999.

[2] H Kang. *Digital Color Halftoning*. SPIE, Bellingham, 1999.

[3] M. Kriss. Modeling an electro-photographic printer: Part i. monochrome systems. In *SPIE Proc.*, volume 6493, pages N1–N11, 2007.

[4] M. Kriss. Modeling an electro-photographic printer: Part ii. color systems. In *SPIE Proc.*, volume 6493, pages O1–O10, 2007.

[5] G. Rodgers. Optical dot gain in halftone print. In R. Eschbach, editor, *Recent Progress in Digital Halftoning II*. IS&T, 1999.

Biography

Dr. Kriss received his BA, MS and PhD in Physics from the University of California at Los Angeles. He joined the Eastman Kodak Research Laboratories in 1969 and until his retirement in 1993 worked on both conventional and digital imaging systems with the emphasis on analytical and computer models. During his tenure at Kodak he spent three years in Japan and headed up the Imaging Processing Laboratory and Algorithm Developing Laboratory. He joined the University of Rochester in 1993 where he was the executive director of the Center for Electronic Imaging Systems and taught through the Computer and Electrical Engineering Department. He joined Sharp Laboratories of America in 2000 where he headed the Color Imaging Group. Dr Kriss retired in 2004 but is still active as a consultant, Adjunct Professor at Portland State University, IS&T activities, and as the Editor in Chief of the Wiley-IS&T Series on Imaging Science and Technology. Dr. Kriss was recently elected president of the International Congress of Imaging Science.

Colour in 3D

Lindsay W. MacDonald
Photogrammetry, 3D Imaging and Metrology Research Group
Department of Civil, Environmental and Geomatic Engineering
University College London
London WC1E 6BT
lindsay.macdonald@ucl.ac.uk

Abstract

Traditional colorimetry is based almost entirely on two-dimensional stimuli, seen in well-controlled viewing environments. Media such as prints on paper, projection in a conference room or cinema, and computer displays fit this paradigm very well because each has a well-defined white point and gamut of the colorants and an isotropic surface gloss.

The external world does not so easily conform to these norms. New materials such as goniochromatic paints and inkjet printing of metallic and variable gloss layers require additional characterisation of the surface BRDF. More significantly, in any real scene, either outdoors or indoors, there may be many different whites, any one of which could be taken as a reference. Solving these problems will be essential for successful augmented reality displays, where computer-rendered effects need to be integrated indistinguishably into the visual scene.

Computer-Aided Reclamation of Lost Art

A New Concept

Maria-Lena Demetriou, Jon Yngve Hardeberg, Gabriel Adelmann
Gjøvik University College
Gjøvik, Norway
marialena.dem@gmail.com,jon.hardeberg@hig.no,adel_g@yahoo.com

Abstract

In this work, the term art reclamation (REC) is proposed. Aim is to use object imaging to reclaim the original form that is irretrievably lost. The basic concept is that of using pre-WW II low-quality colour and B/W photographs of lost paintings on one hand, and contemporary high-quality photographs of surviving stylistically similar paintings by the same master on the other, to attempt an enhancement of the former to an acceptable quality standard.

1 Introduction

Preserving the artistic heritage for present and future audiences has been traditionally dealt with by the sciences of **conservation (C)** and **restoration (R)**, aimed, respectively, at arresting and reversing the action of such factors as ageing of the materials, dirt accumulation, or mechanical, biological, or chemical injuries to the work of art. These time- and skill-intensive technologies have been revolutionized by computer-based methodologies, but remain essentially operator-dependent. Art object imaging plays an ancillary, albeit very important role, for purposes such as documentation, teaching, and mass reproduction.

A self-evident prerequisite of any C/R endeavour is the survival of that work of art, either in its entirety, or in a sufficient 'critical mass' to allow appreciation of the creator's artistic intention and skill.

1.1 Aim

The current project aims at exploring an apparent "lost cause" - that of heritage preservation as applied to paintings documented to have been destroyed at times of major turmoil, such as WW II. The question at hand is not whether these objects can ever be recuperated (clearly, they cannot), but rather, whether it is possible to recuperate, by computer manipulation of existing photographs, a reasonably accurate image of the lost original.

We propose, for this novel question, the term, art **reclamation (REC)**. In this context, art object imaging plays a central role, since it is, literally, an image that is reclaimed, the original being irretrievably lost.

Our reasoning is that if the answer turns out to be affirmative, a high-resolution electronic sample and a corresponding hard-copy reproduction can be obtained and publicly displayed, to the interest and delight of a potentially very large audience, encompassing all age and interest groups, whether technically-, artistically-, or historically-minded. This would represent the sole possibility of contact with lost creations belonging to the core heritage of humanity.

Table 1 circumscribes two main points of difference between REC and P/R.

Table 1: REC vs. P/R

	REC	P/R
Object still existent	No	Yes
Methodology	Acting on existing images of the object and/or of similar objects (imaging essential)	Acting on the object itself (imaging ancillary)

1.2 Methodological underpinnings

The basic concept is that of using pre-WW II low-quality colour and B/W photographs of lost paintings on one hand, and contemporary high-quality photographs of surviving stylistically similar paintings by the same master (in our attempt, P.P. Rubens) on the other, to attempt an enhancement of the former to an acceptable quality standard.

This unique idea has evolved from the interaction of people with different backgrounds, sharing an equal interest in the two core manifestations of the human spirit - art and technology. Traditionally viewed as opposites, they appear to this team as inextricably related.

As usual with a pioneering attempt, it is important to underscore not only what our project attempts to be, but also how it departs from similar endeavours - or, in other words, what it is ***not***. Emphatically, then, this is ***not*** a colorization technology, as we attempt not to ***recreate***, but to ***redeem (reclaim)*** colour-related information, in a manner as objective, operator-independent, and art-historically accurate as possible.

Investigating a pioneering concept, we are in uncharted territory, and have no previous body of research to draw, expand, or improve upon. We harbour no other expectation than that of an objective assessment as to whether our goal is achievable or not, at an adequate level of quality.

2 Proposed Method

At the beginning of the project, after discussion among all the members of the project team, two main problems to be solved were identified. The first one was the essential enhancement of resolution, as the scanned hard copies were both of low quality and had been produced using halftoning. The second problem was the mapping of colour spaces between the low and high quality images.

The approaches to these two problems were chosen as example-based super-resolution and colour correction respectively. In this paper, we focus on example-based super-resolution.

2.1 Example-Based Super Resolution

Conventional super resolution (SR) is performed using a number of measured low-quality images to produce a single higher optical resolution sample. SR recovers the high frequencies in the resulting image by exploiting the given images, where each has a different aliasing effect. Using the measurements in aliased form, the high frequencies which represent details in the final image can be recovered.

However, in this project there are not many low quality images available to be used as given, but instead only a single image is. While the available images are in both colour and greyscale, the approach using various given images is not feasible.

As the information that would be retrieved from many given images is not available, an alternative source of information is needed. An a priori knowledge about the objects should be gained. This is where example-based approach in SR comes in frame. In this method of SR, well-chosen examples are used to apply the regularization effectively.

The application of SR depends strongly on the type of reconstruction therefore it was not confirmed that the project would be feasible. Application of the regularization could even be impossible.

So far SR implementations using a single given image have been done for very specific applications. These approaches were studied and analysed, before selecting the method to be applied.

3 Proposed Super-Resolution Approach

One of the approaches studied was the one implemented by Dmitry Datsenko as described in [8]. It was applied to examples from narrow family of images, specifically on scanned documents and face portraits.

This approach discusses that to choose the regularisation properly, it is best to work with the probability density function (PDF). This way the proper behaviour of the desired image is taken into account and naturally the outcome of the process is of better quality. Next topic in question is of course how to set the PDF. Instead of defining it arbitrarily and intuitively by forcing a simple expression to describe it, image examples are employed to define it.

There are three ways to use examples in such inverse problems: use them to fine-tune the parameters of previously defined regularization expressions [15, 18, 9, 12, 16, 1, 2, 7], use them directly for the reconstruction procedure [6, 17, 11, 10, 14, 4] or fuse the above two techniques somehow [11, 10, 3, 5].

The first method as described above is titled as *learning prior parameters*. This method is not applicable in this project, as its use is to further improve the results if they already reach a sufficient level of quality. [18] If instead the image is not learned prior, but the examples are directly used to learn the posterior PDF which is then used for the reconstruction, we talk about *learning the posterior directly*. Finally, fusing both techniques together covers the approach of *building a regularization expression with examples*. In this case, examples are found as part of the reconstruction process, and then plugged directly into an explicit regularization expression.

The following sections cover the application of learning the posterior directly.

3.1 Examples Database

Examples are gathered in a database of similar structure as Haber's method and from there they are used directly in the reconstruction algorithm.

To obtain a database that is tightly coupled to the degradation that is meant to be restored, i.e. the inverse problem to be resolved, two sets of images are gathered. The first set is that of high quality images, while the second set is that of the corresponding low quality images.

Use of the database is simply pattern matching. Given a low quality image, a look-up in the database is performed to find similar low quality examples. As the corresponding high-quality pair is available, it can be used so that high quality is provided.

Of course, the above approach is not applicable to large images, as high level of similarity is very difficult to be identified in large images. Therefore, the above method is applicable to small patches of images. Typically, these patches have size of 5x5 up to 25x25 pixels. As the size of the patches is very small, the process is applied locally.

This means that it is necessary to go over the low quality images set and extract all image patches of the defined size, while allowing overlapping. This provides a very large set of examples.

Depending on the defined size n of the low quality patch, which will correspond to a high quality patch of different size, the quality of the transform will be different. If the size of the low quality patch is chosen to be small, then many irrelevant patches might be selected, leading to incorrect results. Additionally, the choice of the high quality patch affects strongly, as being large causes relying on the spatial context, while being small wastes information. Size of high quality patch m in [13, 6, 17, 11, 10, 4] is much smaller. According to [5], best results are obtained when using the critical value of m.

3.2 Reconstruction Process

When the database has been established, the reconstruction process can begin. Reconstruction is performed as follows:

For every location [i, j] in the low quality image y, a patch n x n is extracted. The best match for that patch, i.e. the nearest neighbour from the available set of low quality patches Y in the database. To identify the nearest neighbour, a threshold T of the distance from y[i,j] is defined. Threshold T depends on the patch size and the noise variation. All the patches in Y that have smaller or equal difference from y[i,j] are considered. This is mathematically defined as

$$\| y_{[i,j]} - y_k \|_2^2 \leq T \tag{1}$$

Naturally, when the candidate patch or patches have been selected, their corresponding high quality pairs of set X in the database are the ones that will be used in the reconstruction process.

Due to the overlapping in the image search over every [i, j] in the image, a candidate patch x of size $m \times m$ exists. An output canvas $\hat{x}$ can be defined, where all the pixel values can be filled in. Every example x[i,j] found has a known footprint on the canvas. Several methods can be followed to use the resulting matches:

Scalar MMSE Estimate : This method uses all the patches in set X[i,j] that overlap the pixel [I, J]. An approximate MMSE estimate is applied by averaging all the values. By creating a histogram of the values, a 1D approximation description of the posterior p(x—y) is made and the expected value can be computed by a simple mean of the samples. This method is vulnerable to outliers.

Scalar MAP Estimate : Again with the help of the histogram of the values, and by using its peak as the MAP estimation for the desired output. This approach also has its disadvantages, as histogram can be too poor, making curve fitting or smoothing needed.

Non-overlap and 1-NN : Using this algorithm, only the nearest neighbour is selected and if there is no overlapping in the patches used, there is only one value for each [I, J] location. This is the coinciding of the above two methods, as the output at each location is the candidate value. Since this reconstruction process is performed separately for each location [I, J] in the image, it can be defined as a pixel-based reconstruction. As such, the process is very easy to implement, while at the same time, simplicity makes it vulnerable to outliers.

3.3 Building a regularisation expression with examples

As defined above, this approach fuses both of the other two techniques together. Of course, it has advantages and drawbacks.

Firstly, considering an image as a whole and not locally is better. Secondly, joining the measurements and the regularisation together in the likelihood term is a clear definition of the reconstruction objective. However, treating an image locally enables parallelization and simplification of the algorithms. Not only that but also the examples are directly used in the reconstruction process, rather than being based on an expression.

In order to have all advantages, one must apply both techniques. Firstly apply the operation locally, but then, instead of a simple operation directly with the selected patches, one can use a more complicated regularisation expression, in which examples have been fed. Such an approach will have a tighter relation to the final target and will also take into account the whole image. [11, 10, 3, 5] Moreover, this combination of methods into one handles outliers better.

As the expression depends on the examples, there is a strong dependency on the measurements, as they will lead to the selection of the high quality nearest neighbour. [11, 10, 3]

Finally, the approach described in [5] was inspired by the work of Baker and Kanade in [3], as well as the work of Freeman in [11, 10], which apply the above method, with however different expressions.

3.4 Penalties

In this approach, there is also a strategy followed to reduce the number of outliers that interfere with the selected patches. A global penalty function and a MAP penalty function are implemented to choose and cut off the problematic patches.

Weights are given to the examples, so that the examples that fit better with the surroundings are weighted higher, while the poorer ones are underweighted. The examples are sequentially pruned, one patch at a time. The example to be pruned is the one that gives the largest difference between a reference penalty value and the modified MAP penalty value.

A side advantage of this procedure is that a sequence of output images is obtained, following the pruning of the examples.

4 Summary & Further Work

In this paper, the issues regarding the implementation have been discussed, as well as the proposed solution to the matter of super-resolution.

Next step of the project is the comparison between the performance of the algorithm when varying the inner processes. Moreover, paintings that exist in high quality will be inserted into the program and be tested in order to provide a measure of performance.

References

[1] M. Aharon, M. Elad, and A.M. Bruckstein. The k-svd: an algorithm for designing of overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Proc.*, 54(11):4311–4322, 2006.

[2] M. Aharon, M. Elad, and A.M. Bruckstein. On the uniqueness of overcomplete dictionaries, and a practical way to retrieve them. *J. Linear Algebr. Appl.*, 416(11):48–67, 2006.

[3] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(9):1167–1183, 2002.

[4] A. Criminisi, P. Perez, and K. Toyama. Region filling and object removal by exemplar-based image inpainting. *IEEE Trans. Image Process.*, 13(9):1200–1212, 2004.

[5] D. Datsenko and M. Elad. Example-based single document image super-resolution: a global map approach with outlier rejection. *J. Math. Signal Process.*, 2007.

[6] A.A. Efros and T.K. Leung. Texture synthesis by non-parametric sampling. *IEEE Int. Conf. Computer Vision, Corfu, Greece*, pages 1033–1038, 1999.

[7] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.*, 15(12):3736–3745, 2006.

[8] M. Elad and D. Datsenko. Example-based regularization deployed to super-resolution reconstruction of a single image. *The Computer Journal Advance Access, Oxford University Press*, 2007.

[9] K. Engan, S.O. Aase, and J.H. Hakon-Husoy. Method of optimal directions for frame design. *IEEE Int. Conf. Acoustics, Speech, and Signal Processing, Phoenix, Arizona*, 5:2443–2446, 1999.

[10] W.T. Freeman, T.R. Jones, and E.C. Pasztor. Example-based super-resolution. *IEEE Comput. Graphi. Appl.*, 22(2):56–65, 2002.

[11] W.T. Freeman, E.C. Pasztor, and O.T. Carmichael. Learning low-level vision. *Int. J. Comput. Vision*, 40(1):25–47, 2000.

[12] E. Haber and L. Tenorio. Learning regularization functional. *Inverse Problems*, 19:611–626, 2003.

[13] R.L. Lagendijk and J. Biemond. Iterative identification and restoration of images. *Kluwer Academic Publishers, Boston*, 1991.

[14] R. Nakagaki and A.K. Katsaggelos. Vq-based blind image restoration algorithm. *IEEE Trans. Image Process.*, 12(9):1044–1053, 2003.

[15] B.A. Olshausen and D.J. Field. Sparse coding with an overcomplete basis set: a strategy employed by v1? *Vision Res.*, 37:311–325, 1997.

[16] S. Roth and M.J. Black. Fields of experts: a framework for learning image priors. *IEEE Conf. Computer Vision and Pattern Recognition, San-Diego, California*, pages 860–867, 2005.

[17] L.-Y. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. *Proc. of SIGGRAPH, New Orleans, Louisiana*, pages 479–488, 2000.

[18] S.C. Zhu and D. Mumford. Prior learning and gibbs reaction-diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(11):1236–1250, 1997.

Index of Authors