A Mathematical Approach to Best Luminance Maps

ALI ALSAM[1] AND HANS JAKOB RIVERTZ[1]

[1] *Department of Computer Science, Norwegian University of Science and Technology*

Compiled January 29, 2018

An algorithm to calculate the best global mapping from colour to greyscale is presented. We assert that the best mapping minimises the difference between the multi-channel local tensor and the tensor resultant mono-chromatic image. To minimise the objective function, we represent the grayscale image as a weighted sum of the RGB channels, three channels and their second order polynomial and three channels and their root polynomial. The optimisation searches for the best weights to combine the linear, polynomial and root polynomial functions. Our result show that the optimal weights can half the root mean square difference between the colour gradients and those achieved by the conventional lumiance transformation. Further improvement are achieved by adding the squared and root squared channels to the solution. The improvements are also visually evident.

© 2018 Optical Society of America

OCIS codes: (100.0100) Image processing (100.2000) Digital image processing (330.1690) Color, Colour to Gray scale

http://dx.doi.org/10.1364/ao.XX.XXXXXX

1. INTRODUCTION

In a well lit environment, people with normal, trichromatic, colour vision perceive the world around them as being colourful. The advantage of colour over mono-chromatic vision is the added dimensions which allows us to discern a red apple from green leaves even when they share the same luminance or average reflected spectral intensity. One of the challenging problems in colour science, is that of converting a colour image to grey-scale in a manner that preserves the distinctions between the various colour regions and objects.

The conversion of higher dimensional images to grey-scale is important in a wide range of applications such as: When printing a colour image on a mono-chromatic printer, when viewing multi-spectral images on colour monitors and when we wish to create a pleasing wedding photo. While the exact definition of a successful conversion might be defined by the actual problem such as detecting environmental changes from satellite images or converting a modern child's photo to resemble an old photograph there is almost always a general aim that is to preserve the visual information which allow us to make sense of information.

A very general definition of visible contrast is difference between various parts of the scene or image. A high contrast grayscale image has a maximum value that represents white and a minimum that represents black while a low contrast image has a maximum and minimum that are visually close such as two levels of mid grey. In the colour space, on the other hand, contrast is encoded along red-green, blue-yellow and white-black thus the problem of converting colour images to an optimised grey-scale representation might be stated as the problem of encoding, blue-yellow, red-green contrast in the single available dimension namely: white to black.

Today, there is a vast literature addressing colour to grey conversion and image fusion with the shared focus of preserving colour contrast and the discernibility of different image regions. Generally speaking, the literature can be divided into two main classes: in the first, there are global methods [1] [2] and in the second, there are local spatial algorithms [3] [4], [5] [6] that assign a grey-value to each pixel based on its difference from neighbouring pixels.

The algorithm presented in this article addresses a mathematical question, namely: If, in a trichromatic vision system, only scaling of the channels was allowed, what would the best luminance direction be?

Although many studies show that human colour perception is influenced by many non-physical factors including a person's natural language [7, 8], it is practically assumed that a global conversion from colour to gray is a weighted linear sum of the colour channels where the weight assigned to green is greater than that assigned to red and blue.

In this paper we adhere to the approach of the weighted linear sum and ask: what are the optimal image specific weights that combine the red, green and blue channels? We would like to underline the difference between an image dependent luminance direction and that derived from vision experiments as in the latter the weights are experimentally estimated using patches of colours and observers who indicate which patch is lighter, i.e. the approach does not include the scene's spatial variations.

In today's, image processing packages such as photoshop and lightroom, the user is allowed to adjust the weights used in the conversion until a subjective pleasing result is achieved. From this perspective, the objective of this work is to mathematically estimate the optimal weights where our optimality criterion is that the structure tensor of the resultant grey-scale image is as close as possible to that of the original colour version.

We choose the structure tensor because it encodes the colour gradients of the image. Thus the algorithm is designed to search for weights that minimize the difference between the gradients of the colour image represented by the local tensor and those of the gray image.

2. THEORY

A color image may be represented in some trichromatic colour space such as RGB by $\mathbf{I} = (R, G, B)$. A linear luminance representation of the image is the weighted sum $L = w_1 R + w_2 G + w_3 B$, where w_1, w_2, and w_3 are positive constant weights. These weights are chosen so that the grey scale image is as close as possible to the perceived average luminance of the original colour image. Since the human vision is more sensitive to green light than red and blue, w_2 is greater than w_1 and w_3. The values $(w_1, w_2, w_3) = (0.29, 0.59, 0.12)$ are typical values.

Colours that have different hues and identical luminance are known as iso-luminant and they are common in art and nature and the use of global, fixed weights to convert three channel images to grey-scale can and does, in the worst case scenario, lead to converting adjacent colours that are noticeably different in hue to the same gray value.

Instead of defining the gray-scale image as a linear combination of the three colour channels, we can write a more general global luminance using a second degree polynomial: $L = w_1 R + w_2 G + w_3 B + w_4 GB + w_5 RB + w_6 RG + w_7 R^2 + w_8 G^2 + w_9 B^2$.

We also consider a root polynomial luminance representation. That is $L = w_1 R + w_2 G + w_3 B + w_4 \sqrt{GB} + w_5 \sqrt{RB} + w_6 \sqrt{RG}$. Root polynomials have recently been used to model non linear relations in color science. See Finlayson, [9].

When we adapt the experimental weights used to convert colour to luminance, we can write the structure tensor of the color image as: $M_C = 0.29 \nabla R \nabla R^T + 0.59 \nabla G \nabla G^T + 0.12 \nabla B \nabla B^T$.

The corresponding luminance image L has structure tensor $M_L = \nabla L \nabla L^T$. By using the matrix norm we have $\|M_L - M_C\|^2 = \mathrm{tr}(M_L - M_C)^2$.

In the following mathematical derivation, we describe the linear case with three weights. The polynomial and root polynomial cases are equivalent. We will find weights w_1, w_2, and w_3 that minimizes the function $W(w_1, w_2, w_3) = \int_\Omega \|M_L - M_C\|^2 \, d\omega$.

We use the fact that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ and get

$$\begin{aligned}\|M_L - M_C\|^2 &= \mathrm{tr}(\nabla L \nabla L^T - M_C)^2 \\ &= \mathrm{tr}(\nabla L \nabla L^T \nabla L \nabla L^T) - \mathrm{tr}(\nabla L \nabla L^T M_C) \\ &\quad - \mathrm{tr}(M_C \nabla L \nabla L^T) + \mathrm{tr}\, M_C^2 \\ &= \mathrm{tr}(\nabla L^T \nabla L \nabla L^T \nabla L) - \mathrm{tr}(\nabla L^T M_C \nabla L) \\ &\quad - \mathrm{tr}(\nabla L^T M_C \nabla L) + \mathrm{tr}\, M_C^2 \\ &= \nabla L^T (\nabla L \nabla L^T - 2M_C) \nabla L + \mathrm{tr}\, M_C^2.\end{aligned}$$

The derivative of ∇L with respect to w_1 is ∇R. Therefore, the derivative of $\|M_L - M_C\|^2$ with respect to w_1 is

$$\begin{aligned}(\|M_L - M_C\|^2)_1 &= \nabla R^T (\nabla L \nabla L^T - 2M_C) \nabla L \\ &\quad + \nabla L^T \nabla R \nabla L^T \nabla L + \nabla L^T \nabla L \nabla R^T \nabla L \\ &\quad + \nabla L^T (\nabla L \nabla L^T - 2M_C) \nabla R.\end{aligned}$$

Reorganising by using $\nabla R^T \nabla L = \nabla L^T \nabla R \in \mathbb{R}$ and $M_C^T = M_C$ gives $(\|M_L - M_C\|^2)_1 = 4 \nabla R^T (M_L - M_C) \nabla L$. Therefore, the derivatives of W are

$$W_1(w_1, w_2, w_3) = 4 \int_\Omega \nabla R^T (M_L - M_C) \nabla L \, d\omega,$$

$$W_2(w_1, w_2, w_3) = 4 \int_\Omega \nabla G^T (M_L - M_C) \nabla L \, d\omega,$$

and

$$W_3(w_1, w_2, w_3) = 4 \int_\Omega \nabla B^T (M_L - M_C) \nabla L \, d\omega.$$

These functions are polynomials in w_1, w_2, and w_3. E.g., $W_1(w_1, w_2, w_3) = c_{300} w_1^3 + c_{030} w_2^3 + c_{003} w_3^3 + c_{210} w_1^2 w_2 + c_{201} w_1^2 w_3 + c_{120} w_2^2 w_1 + c_{021} w_2^2 w_3 + c_{102} w_3^2 w_1 + c_{012} w_3^2 w_2 + c_{100} w_1 + c_{010} w_2 + c_{001} w_3$, where $c_{300} = 4 \int_\Omega (\nabla R^T \nabla R)^2 \, d\omega$, and so on.

We solve $\mathbf{F} = (W_1, W_2, W_3) = 0$ by using the multivariate Newton method. The Jacobian $D\mathbf{F}$ is symmetric and has coefficients

$$W_{11}(w_1, w_2, w_3) = \\ 4 \int_\Omega \nabla R^T (\nabla R \nabla L^T + \nabla L \nabla R^T) \nabla L + \nabla R^T (M_L - M_C) \nabla R \, d\omega,$$

$$W_{12}(w_1, w_2, w_3) = \\ 4 \int_\Omega \nabla R^T (\nabla G \nabla L^T + \nabla L \nabla G^T) \nabla L + \nabla R^T (M_L - M_C) \nabla G \, d\omega,$$

$$W_{13}(w_1, w_2, w_3) = \\ 4 \int_\Omega \nabla R^T (\nabla B \nabla L^T + \nabla L \nabla B^T) \nabla L + \nabla R^T (M_L - M_C) \nabla B \, d\omega,$$

$$W_{22}(w_1, w_2, w_3) = \\ 4 \int_\Omega \nabla G^T (\nabla G \nabla L^T + \nabla L \nabla G^T) \nabla L + \nabla G^T (M_L - M_C) \nabla G \, d\omega,$$

$$W_{23}(w_1, w_2, w_3) = \\ 4 \int_\Omega \nabla G^T (\nabla B \nabla L^T + \nabla L \nabla B^T) \nabla L + \nabla G^T (M_L - M_C) \nabla B \, d\omega,$$

and

$$W_{33}(w_1, w_2, w_3) = \\ 4 \int_\Omega \nabla B^T (\nabla B \nabla L^T + \nabla L \nabla B^T) \nabla L + \nabla B^T (M_L - M_C) \nabla B \, d\omega.$$

The code is available by request.

3. RESULTS

Five images are chosen to present the performance of the method: a colourful graphic image 1a, a painting by Monet 2a, an image of 5 different colour caps 3a, a photo of a baboon 4a, and a picture of two parrots 5a. Prior to running the algorithm, a Gaussian blur filter with $\sigma = 3$ was applied to the the caps image and sunset image. Table 1 shows the root mean square error $RMSE = \sqrt{\frac{W}{N}}$ of the greyscale structure tensor relative to the colour structure tensor. Where, N is the number of pixels. We observe that the $RMSE$ improves to approximately one half in

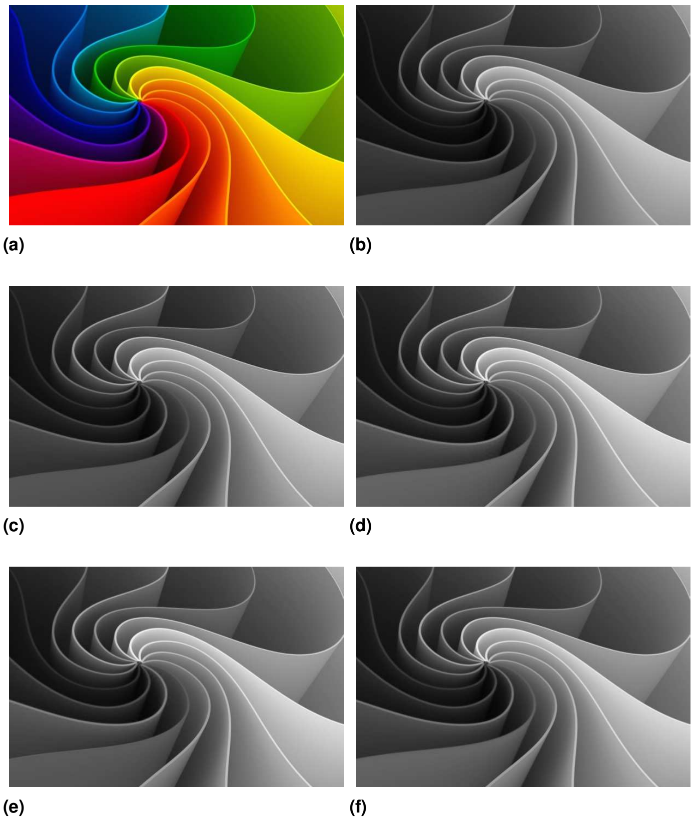


Fig. 1. The figure displays the original colour image 1a, the standard luminance image 1b, the best linear greyscale image 1c, the best second degree root polynomial greyscale image 1d, the best third degree root polynomial image 1e, and the best polynomial greyscale image 1f,

our method compared to the $RMSE$ of the standard luminance image.

This improvement is also evident in the visual inspection of the sample images. As an example, the blue sheet in Figure 1a is more visible in Figure 1c to Figure 1f than it is in the standard luminance image in Figure 1b. The improvement is gentle and does not reduce the contrast in other part of the image. The orange sun in Monets painting in Figure 2a almost disappears in the luminance image in Figure 2b. The sun is, however, clearly visible in Figures 2c-2f as well as many other details such as the reflection in the sea. Similar improvements are visible in the image series of caps in Figure 3. There are minor improvements in the image series in Figure 4 and Figure 5.

Numerically, the algorithm converged to the coefficients shown in table 2, 3, 4, and 5. Although, some the the coefficients in table 5 are relatively large and some are negative, the range of the resulting greyscale images was between 0 and 1. No contrast reducing renormalizing was therefore performed or necessary.

4. DISCUSSION

In this article, we presented an algorithm to map colour images to greyscale with the condition that the mapping should preserve as much as possible of the gradients between the various image parts represented by the tensor matrix.

The algorithm is global, i.e. any two pixels with the same RGB value will map to the same greyscale independent of their location in the image or the colours that surround them.

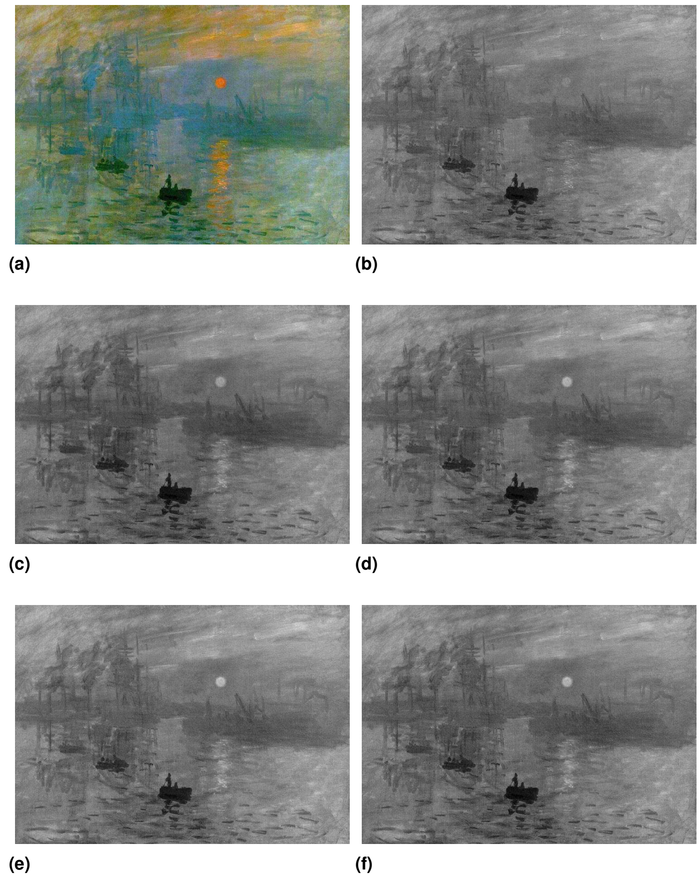


Fig. 2. The figure displays the original colour image 2a, the standard luminance image 2b, the best linear greyscale image 2c, the best second degree root polynomial greyscale image 2d, the best third degree root polynomial image 2e, and the best polynomial greyscale image 2f,

To find the best luminance mapping we represent the problem as a system of n cubic polynomial equations where the number of solutions is at most 3^n. One solution is the trivial solution $(0, 0, 0, \ldots)$ and for any solution, its negative is also a solution. The solutions of the system are critical points and not automatically local minima. A critical point is a local minimum when all n eigenvalues of the Hessian $H = DF$ are positive. That limits the number of local minima to a few.

Although the method is fast and can easily be used in practical image processing, our interest in this work is theoretical where we wanted to demonstrate that a mathematical optimality condition pertaining to preserving the gradients of the colour can be used to result in an improved global colour to greyscale conversion.

Our result show that even when we constrain the mapping to be a linear combination of the colour channels, it is possible to improve upon the root mean square of luminance conversion by a factor of two. Such an improvement is interesting as it can easily be performed to optimise the image contrast by a simple scaling.

	Colour sheets	Caps	Sunset
Luminance	1.21e-03	1.33e-04	4.95e-05
Linear	8.78e-04	1.01e-04	3.66e-05
Root Poly 2. deg.	6.15e-04	8.16e-05	3.18e-05
Root Poly 3. deg.	4.80e-04	8.02e-05	2.94e-05
Poly 2. deg.	5.53e-04	8.30e-05	2.92e-05

Table 1. Root mean square error of the greyscale structure tensor relative to the color structure tensor.

	Sample	Caps	Sunset
R	0.324	0.483	0.508
G	0.569	0.472	0.420
B	0.185	0.0522	0.0251

Table 2. The coefficients for best linear greyscale. The initial vector was $(1, 1, 1)$.

	Sample	Caps	Sunset
R	0.275	0.474	0.638
G	0.615	0.462	0.418
B	0.186	0.0657	-0.104
GB	-0.188	-0.714	0.667
RB	-0.025	0.698	-0.0811
RG	-0.178	-0.560	-0.981
R^2	0.110	0.0213	0.387
G^2	0.118	0.533	0.208
B^2	0.0857	0.0285	-0.156

Table 3. The coefficients for best second degree polynomial greyscale. The initial vector was the best linear greyscale from table 2.

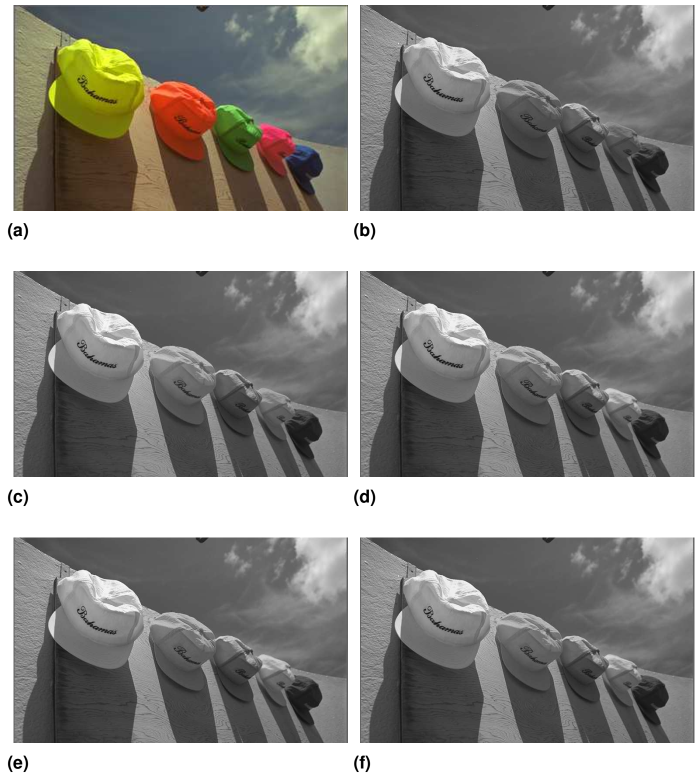
	Sample	Caps	Sunset
R	0.399	0.694	0.844
G	0.760	1.73	0.469
B	0.248	0.0604	0.152
$\sqrt{GB}$	-0.172	-1.13	0.214
$\sqrt{RB}$	0.0388	1.15	-0.400
$\sqrt{RG}$	-0.181	-1.51	-0.301

Table 4. The coefficients for best second degree root polynomial greyscale. The initial vector was the best linear greyscale from table 2.

	Sample	Caps	Sunset
R	0.118	0.462	6.12
G	0.374	-1.97	0.549
B	-0.0049	0.0625	-2.80
$\sqrt{GB}$	-2.77	-13.7	-11.2
$\sqrt{RB}$	-0.539	13.7	-6.10
$\sqrt{RG}$	-4.27	-13.8	36.4
$\sqrt[3]{GB^2}$	1.59	9.72	9.40
$\sqrt[3]{BR^2}$	0.495	-2.79	2.70
$\sqrt[3]{RG^2}$	2.39	12.0	-12.9
$\sqrt[3]{G^2B}$	1.49	7.90	4.37
$\sqrt[3]{B^2R}$	0.366	-9.67	4.94
$\sqrt[3]{R^2G}$	2.25	4.21	-29.3
$\sqrt[3]{RGB}$	-0.419	-5.11	-1.22

Table 5. The coefficients for best third degree root polynomial greyscale. The initial vector for 'Sample' and 'Sunset' was the coefficients of the best second degree root polynomial greyscale from table 4. The initial vector for 'caps' was (1,1,1)

.



(a) (b) (c) (d) (e) (f)

Fig. 3. The figure displays the original colour image 3a, the standard luminance image 3b, the best linear greyscale image 3c, the best second degree root polynomial greyscale image 3d, the best third degree root polynomial image 3e, and the best polynomial greyscale image 3f,

REFERENCES

1. M. Grundland and N. A. Dodgson, "Decolorize: Fast, contrast enhancing, color to grayscale conversion," in "Pattern Recognition, 40(11)," (2007), pp. 2891–2896.
2. D. Connah, G. Finlayson, and M. Bloj, "Seeing beyond luminance: A psychophysical comparison of techniques for converting colour images to greyscale," in "15th Color Imaging Conference: Color, Science, Systems and Applications.", (2007), pp. 336–341.
3. A. Alsam and M. S. Drew, "Fastcolour2grey," 16th Color Imaging Conference: Color, Science, Systems and Applications, Society for Imaging Science & Technology (IS&T)/Society for Information Display (SID) joint conference, Portland, Oregon. pp. 342–346 (2008).
4. R. Bala and R. Eschbach, "Spatial color-to-grayscale transform preserving chrominance edge information," in "14th Color Imaging Conference: Color, Science, Systems and Applications.", (2004), pp. 82–86.
5. L. Neumann, M. Cadík, and A. Nemcsics, "An efficient perception-based adaptive color to gray transformation," in "Computational Aesthetics," , D. W. Cunningham, G. W. Meyer, L. Neumann, A. Dunning, and R. Paricio, eds. (Eurographics Association, 2007), pp. 73–80.
6. K. Rasche, R. Geist, and J. Westall, "Re-coloring images for gamuts of lower dimension," Comput. Graph. Forum **24**, 423–432 (2005).
7. P. Kay and W. Kempton, "What is the sapir-whorf hypothesis?" American anthropologist **86**, 65–79 (1984).
8. K. Zhou, L. Mo, P. Kay, V. P. Kwok, T. N. Ip, and L. H. Tan, "Newly trained lexical categories produce lateralized categorical perception of color," Proceedings of the National Academy of Sciences **107**, 9974–9978 (2010).
9. G. D. Finlayson, M. Mackiewicz, and A. C. Hurlbert, "Color correction using root-polynomial regression," IEEE Trans. Image Processing **24**,

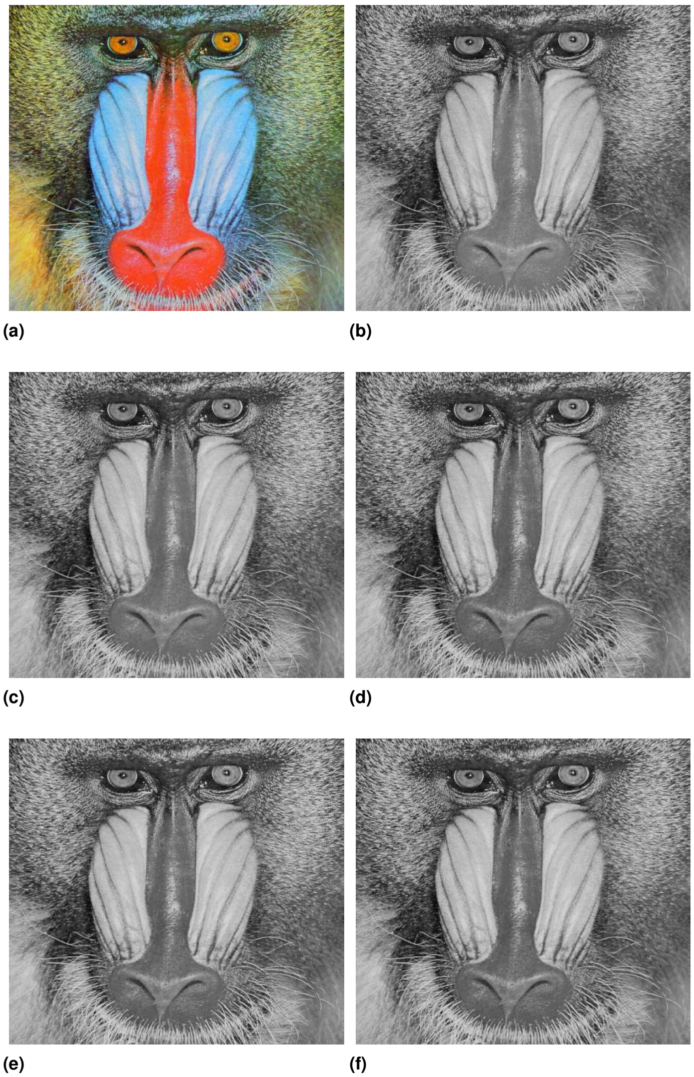


Fig. 4. The figure displays the original colour image 4a, the standard luminance image 4b, the best linear greyscale image 4c, the best second degree root polynomial greyscale image 4d, the best third degree root polynomial image 4e, and the best polynomial greyscale image 4f,

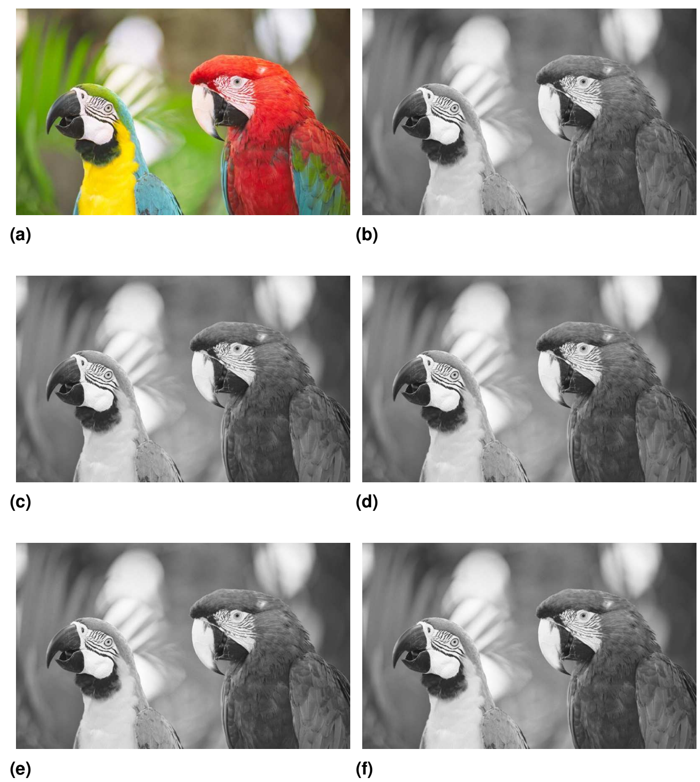


Fig. 5. The figure displays the original colour image 5a, the standard luminance image 5b, the best linear greyscale image 5c, the best second degree root polynomial greyscale image 5d, the best third degree root polynomial image 5e, and the best polynomial greyscale image 5f,

1460–1470 (2015).