

Art_7.07

LINEAR VERSUS NON-LINEAR DIGITAL IMAGE PROCESSING METHODS

João Carlos Aparício Fernandes

Departamento de Electrónica Industrial da Universidade do Minho

Guimarães, Portugal

Email: aparicio@dei.uminho.pt

ABSTRACT

Digital image processing methods can be classified either as linear or as non-linear. The robust mathematical basis for the linear approach justifies the preference for linear methods, although the non-linear techniques are spreading. Analysing the image gathering chain, the non-linearity of the intensity to voltage conversion, known as the gamma function, is referred but its consequences usually disregarded. This may put in question the validity of the straight linear approach this chain is not, and explain why apparent formally perfect techniques often lead to inconclusive or unexpected results. On the other hand, non-linear approaches, as those based on mathematical morphology, are gaining wider preference, mainly due to their simplicity and robustness.

1. INTRODUCTION

Digital images result from an analog to digital conversion of an electric signal from the image sensor. The digital format is convenient for archiving, transmission and processing with complex algorithms not available in the analog world.

The Cathode Ray Tube (CRT) is the reference device for display of electronic images. Some characteristics of CRTs are embedded in the electronic image applications, and digital image processing could not escape. The main point of this paper concerns the use of a non-linear pre-correction of the voltage signal. This is usually referred, but not accounted for (Fernandes, 2004).

This non-linearity in the chain leads to a non-linear process, even when the algorithm that supports the processing is formally linear. The results of a supposed linear technique fail often, just for a small differences in the image, and that should not happen in a linear approach.

In another perspective, basically non-linear algorithms are reported successful, and their use is increasing. Mathematical Morphology (Serra, 1982) is one of these. Its mathematical background and implementation simplicity provides robust applications development.

2. THE GAMMA CORRECTION

In early television development, it was found that the light intensity from the CRT and the voltage signal relate in a non-linear way (Poynton, 1996). The relationship was modelled as a power function in the form $y = x^\gamma$, where γ , the exponent of the power function was a

constant with a practical value of 2.5. From the name of the Greek character used resulted the usual name – the gamma function.

To account for this effect, it was decided to pre-correct the electric signal at the source by the inverse function. In this way, the full chain result is $i = \left(u^{\frac{1}{\gamma}}\right) = u$, producing a linear

relationship between the intensity on the image sensor and the intensity on the CRT faceplate image. Doing the pre-correction at the image acquisition level allowed for a single calibration dependent equipment and simpler receiving systems, just using the natural response of the CRTs. In this way, the video signal is pre-compensated at the source in a non-linear way, and the image is perfectly reproduced even using a complex chain of transmission, as far as it is kept linear. Practice recommended to compensate only for a gamma=2.2 (Poynton, 1996), producing a marginally more pleasant image.

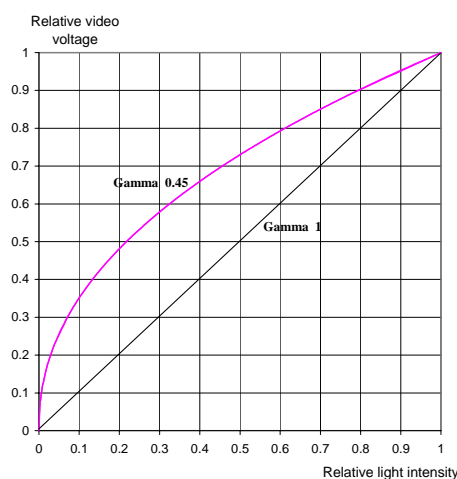


Figure 1 – The inverse gamma correction

Introducing the analog to digital (AD) and digital to analog (DA) conversions to the electric signal, the system functions as expected, as far as linear relationships are kept.

The hidden problem may appear when using the numeric values in the digital image to account for the image intensity. The values of the digital image are a linear translation of the video signal, and relate to light intensities by the inverse-gamma function, a non-linear relationship. For example, the usual average calculation produces the average of the voltage values, and not an average light intensity as intended. Most often, the errors are small and pass unperceived; only processes insensitive to monotonic transformations of intensity (Haralick, 1973) can be guaranteed to work as intended.

Obviously, this problem can be taken into account in the digital domain, providing the gamma and inverse gamma functions in look-up tables (LUTs). But then, another problem arises: if 8 bits are enough for practical purposes in the non-linear representation, a more complex 16 bit is needed for an equivalent linear coding. Because the power function is a close match to a logarithmic transform, the simple 8 bit produce a sufficient signal to noise ratio (SNR) that would be not enough in a linear coding, as can be seen in Table 1. The SNR for the lowest non-zero value ($i=1$) is 0 dB, while the corresponding value for the inverse gamma corrected value is 26.25 dB, corresponding to an equivalent intensity of 20.5.

To guarantee the same SNR for a linear coding, the minimum value should be at least 20, and for a linear scale, the maximum needed to be 20.5×255 , that is 5225, needing 13 bits. For practical reasons, values are stored in a double byte (16 bits) format, needing twice the storage space.

This is the main reason why the simple 8 bit structure with an inverse gamma correction for intensity is the standard intensity coding, and not the double byte linear code, even for the new display techniques as the Liquid Crystal Displays (LCD) that imply a totally different approach to the intensity control.

Table 1 – Inverse Gamma correction

Linear			Gamma		
n	n/255	SNR		i	SNR
0	0		0	0	
1	0.0039	0.00	0.08056	20.54274	26.25
2	0.0078	6.02	0.11040	28.15076	28.99
3	0.0118	9.54	0.13274	33.84789	30.59
4	0.0157	12.04	0.15128	38.57642	31.73
5	0.0196	13.98	0.16743	42.6945	32.61
6	0.0235	15.56	0.18190	46.38349	33.33
7	0.0275	16.90	0.19510	49.75008	33.94
8	0.0314	18.06	0.20731	52.86323	34.46
9	0.0353	19.08	0.21871	55.77054	34.93
10	0.0392	20.00	0.22944	58.50645	35.34
20	0.0784	26.02	0.31441	80.17437	38.08
30	0.1176	29.54	0.37804	96.4	39.68
40	0.1569	32.04	0.43085	109.867	40.82
50	0.1961	33.98	0.47684	121.5955	41.70
60	0.2353	35.56	0.51805	132.1018	42.42
80	0.3137	38.06	0.59042	150.5564	43.55
100	0.3922	40.00	0.65344	166.6285	44.43
120	0.4706	41.58	0.70991	181.0259	45.15
140	0.549	42.92	0.76143	194.165	45.76
160	0.6275	44.08	0.80908	206.3151	46.29
180	0.7059	45.11	0.85358	217.6618	46.76
200	0.7843	46.02	0.89545	228.3395	47.17
220	0.8627	46.85	0.93509	238.4492	47.55
240	0.9412	47.60	0.97282	248.069	47.89
250	0.9804	47.96	0.99104	252.715	48.05
255	1	48.13	1	255	48.13

3. THE LINEAR DILEMMA

Using digital image data without gamma compensation, leads to algorithms with hidden non-linear characteristics, lose their intended purpose and producing distorted results.

If the algorithm uses pure numerical data, and so unrelated to image intensity, the linearity of the process is guaranteed, but the results cannot be related linearly to the image information (that is intensity) nor translated into an image of the same type.

When using linear algorithms for image processing, the input must be a linear representation of image intensities, and so gamma corrected, and with sufficient resolution not to impair the SNR of the representation, that is, a 16-bit model. For an 8-bit starting image, a gamma corrected 16-bit image should be produced, processed, and finally inverse-gamma transformed to another 8-bit output image, to be displayed in the normal gamma dependent system.

The choice of non-linear algorithms, as in the morphological based approach may be viable alternatives. The erosion/dilation and median transformations – if correctly implemented, which is not always the case – are insensitive to monotonic transformations; however, the Top-Hat transform, that uses image intensity differences, needs a very careful implementation.

4. CONCLUSIONS

Linear algorithms in image processing may fail to produce the envisaged results but, instead of simply abandoning the approach, using correctly linear coded images may validate the technique. Probably, many perfectly valid schemes were discarded just for lack of results due to the implicit non-linear chain in the image acquisition equipment.

REFERENCES

1. Fernandes, J.C.A., *Análise de Imagens Texturadas com base em Matrizes de Co-ocorrência Modificadas*, 2004, Universidade do Minho, Guimarães.
2. Haralick, R.M., K. Shanmugam, and I. Dinstein, *Texture features for image classification*, IEEE Trans. Systems Man Cybernet., 1973, **SMC-3**: p. 610-621.
3. Poynton, C., *A Technical Introduction to Digital Video*, 1996: John Wiley & Sons. 321 pages.
4. Serra, J., *Image Analysis and Mathematical Morphology*, vol. 1, 1982: Academic Press.