

IMAGE QUALITY IMPROVEMENT BY ADAPTIVE EXPOSURE CORRECTION TECHNIQUES

Messina G., Castorina A., Battiato S., Bosco A.,
{giuseppe.messina; alfio.castorina; sebastiano.battiato; angelo.bosco}@st.com.

STMicroelectronics - Str. Primosole 50, 95121 Catania - Italy

ABSTRACT

The proposed paper concerns the processing of images in digital format and, more specifically, particular techniques that can be advantageously used in digital still cameras for improving the quality of images acquired with a non-optimal exposure. The proposed approach analyses the CCD/CMOS sensor Bayer data or the corresponding color generated image and, after identifying specific features, it adjusts the exposure level according to a 'camera response' like function.

1. INTRODUCTION

Digital still cameras are currently among the devices most commonly employed for acquiring digital images [5]. Sensors of ever greater resolution having low-cost/low consumption digital signal processors are readily available in commerce; this has led to the development of digital still cameras capable of acquiring images of very considerable resolution and quality. Image improvement is obtained by increasing the resolution of the sensor and/or by using more sophisticated image-processing algorithms [1], [3], [6] and [8].

A difficult problem to solve is represented by estimating the best exposure to be used in the acquisition phase. This is particularly true for handset devices (e.g. mobile phones) where several factors contribute to acquire bad-exposed pictures: poor optics, absence of flashgun, etc.

There is not an exact definition of what a correct exposure should be. It is possible to abstract a generalization and to define as best the particular exposure that enables one to reproduce the most important regions (according to contextual or perceptive criteria) with a level of grey, or brightness, more or less in the middle of the possible range.

Our method differs from the one described in [4] since the whole process can also be implemented directly on Bayer pattern images [2]. Simpler statistical measures (i.e. focus and contrast) are used to identify information carrying regions. Furthermore, a new features-extraction method based on skin detection is introduced. The

selection of image areas having particular characteristics allows to apply selective exposure correction. These exposure adjustment techniques are designed essentially for mobile sensors applications.

The document is organized as follows. Next section describes both the features-extraction (i.e. focus, contrast and skin) and the exposure correction technique used for automatic enhancement. In section 3 experimental results show the effectiveness of the proposed technique. A conclusion Section closes the paper tracking directions for future work.

2. APPROACH DESCRIPTION

The proposed Exposure Correction algorithm is defined as follows:

- Luminance extraction. If the algorithm is applied on Bayer data, in place of the Y plane that is used in color images, the original frame is sub-sampled by considering the Green color channel only;
- After luminance extraction, some image features are analyzed in order to identify the regions that contain more information. If dermis is detected, a skin extraction algorithm is used; otherwise a contrast and focus features-extraction is adopted;
- Once the 'visually important' pixels were identified (e.g. the pixels belonging to skin features) a global tone correction technique is applied by using the mean gray levels of identified regions as main parameter.

2.1. Contrast and Focus Features

In order to identify regions of the image that contain more information the luminance plane is subdivided in N blocks of equal dimensions (in our experiments N=16). For each block, statistical measures of "contrast" and "focus" are computed. Therefore it is assumed that well focused or high-contrast blocks are more relevant compared to the others. Contrast refers to the range of tones present in the image. A high contrast leads to a higher number of clustered pixels inside a block.

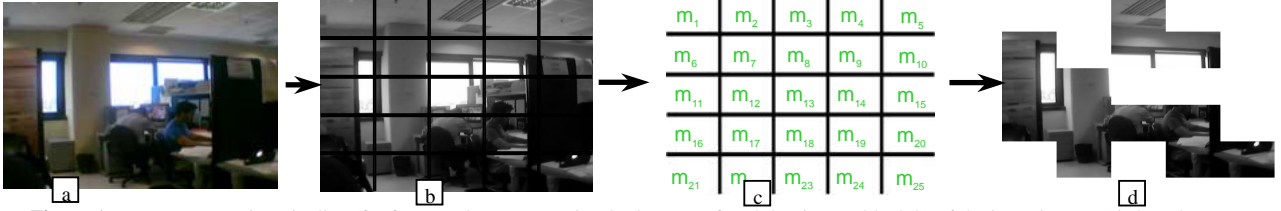


Figure 1. Features extraction pipeline (for focus and contrast). Visual relevance of each luminance block b) of the input image a) is based on relevance measures c) able to obtain a list of relevant blocks d).

Focus characterizes the sharpness or edginess of the block and is useful in identifying regions where high frequency components (i.e. details) are present.

If the aforementioned measures were simply computed on highly underexposed images, then the regions having better exposure would always have higher contrast and edginess compared to those that are obscured. In order to perform a visual analysis revealing the most important features regardless to lighting conditions, a new ‘visibility’ image’ is constructed by pushing the mean grey level of the input green Bayer pattern plane to 128. The push operation is performed using the same function that is used to adjust the exposure level and it will be described later.

The contrast measure is computed by simply building an histogram for each block and then calculating its deviation (2) from the mean value (3). A high deviation value denotes good contrast and vice versa. In order to remove irrelevant peaks, the histogram is slightly smoothed by replacing each entry with its mean in a neighborhood of ray 2. Thus, the original histogram entry $I[i]$ is replaced with the grey-level $\tilde{T}[i]$:

$$\tilde{T}[i] = \frac{(I[i-2] + I[i-1] + I[i] + I[i+1] + I[i+2])}{5} \quad (1)$$

Histogram deviation D is computed as :

$$D = \frac{\sum_{i=0}^{255} |i - M| \cdot \tilde{T}[i]}{\sum_{i=0}^{255} \tilde{T}[i]} \quad (2)$$

where M is the mean value :

$$M = \frac{\sum_{i=0}^{255} i \cdot \tilde{T}[i]}{\sum_{i=0}^{255} \tilde{T}[i]} \quad (3)$$

The focus measure is computed by convolving each block with a simple 3x3 laplacian filter.

In order to discard irrelevant high frequency pixels (mostly noise), the outputs of the convolution at each pixel are thresholded. The mean focus value of each block is computed as:

$$F = \frac{\sum_{i=1}^N \text{thresh}[\text{lapl}(i), T]}{N} \quad (4)$$

where N is the number of pixels and the $\text{thresh}()$ operator discards values lower than a fixed threshold T . Once the values F and D are computed for all blocks, relevant regions will be classified using a linear combination of both values. Features extraction pipeline is illustrated in Fig.1.

2.2. Skin Recognition

Most existing methods for skin colour detection usually threshold a skin similarity measure for each pixel. Human skin colours form a special category of colours, distinctive from the colours of the majority of other natural objects. It has been found that human skin colours are clustered in various colour spaces [9], [11] and [13]. The skin colour variations between people are mostly due to intensity differences. These variations can therefore be reduced using chrominance components only.

Yang et al [12] have demonstrated that the distribution of human skin colours can be represented by a two-dimensional Gaussian function on the chrominance plane. The center of this distribution is determined by the mean vector $\vec{\mu}$ and its shape is determined by the covariance matrix Σ ; both values can be estimated from an appropriate training data set. The conditional probability $p(\vec{x}/s)$ of a block belonging to the skin colour class, given its chrominance vector \vec{x} is then represented by:

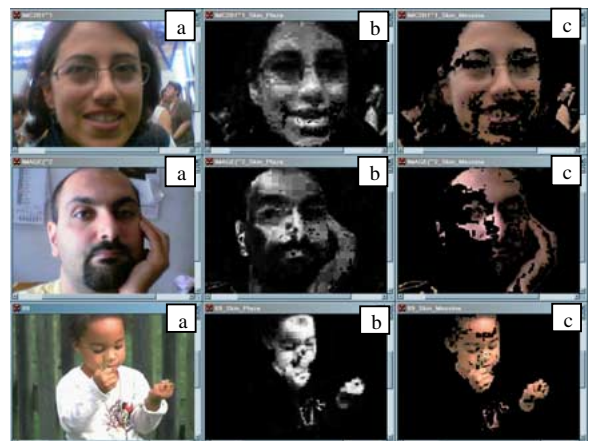


Figure 2. Skin recognition examples on RGB images. a) Original images acquired by Nokia 7650 phone (first and second row) with VGA sensor and compressed in Jpeg format; b) Probabilistic Threshold Output; c) Simplest Threshold Method Output. Third rows image (a) is a standard test image.

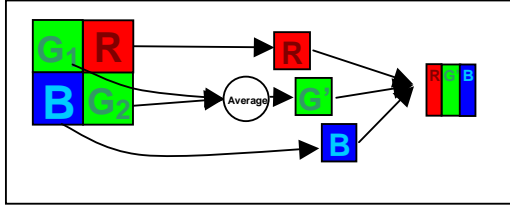


Figure 3. Bayer data sub-sampling generation.

$$p(\vec{x}/s) = \frac{1}{2\pi} \left| \sum \right|^{\frac{1}{2}} \exp \left\{ -\frac{[d(\vec{x})]^2}{2} \right\} \quad (5)$$

where $d(\vec{x})$ is the so-called Mahalanobis distance from the vector \vec{x} to the mean vector $\vec{\mu}$ and defined as:

$$[d(\vec{x})]^2 = (\vec{x} - \vec{\mu})' \Sigma^{-1} (\vec{x} - \vec{\mu}) \quad (6)$$

The value $d(\vec{x})$ determines the probability that a given block belongs to the skin colour class. The larger the distance $d(\vec{x})$, the lower the probability that the block belongs to the skin colour class.

Due to the large quantity of colour spaces, distance measures, and two-dimensional distributions, many skin recognition algorithms can be used. The proposed skin colour algorithm is independent from exposure correction, thus, depending on the color space used (YCrCb or RGB), we introduce two different alternative techniques aimed to recognize skin regions:

1. By using the input YCbCr image and the conditional probability (5), each pixel is classified as belonging to a skin region or not. Then a new image with normalized grayscale values is derived, where skin areas are properly highlighted (Fig.2.b). The higher the grey value the bigger the probability to compute a reliable identification.
2. By processing an input RGB image, a 2D chrominance distribution histogram (r, g) is computed, where $r=R/(R+G+B)$ and $g=G/(R+G+B)$. Chrominance values representing skin are clustered in a specific area of the (r, g) plane, called “*skin locus*”, as defined in [11]. Pixels having a chrominance value belonging to the skin locus will be used to correct exposure (Fig.2.c).

The conversion from Bayer data to YCrCb is more expensive than a conversion to RGB; furthermore, we are interested only in determining the areas in which skin pixels are located. Thus we prefer to perform a light RGB color interpolation by sub-sampling the original Bayer data, as illustrated in Fig.3. Each pixel of the quarter-sized RGB color image corresponds to a group of four pixels in the Bayer domain.

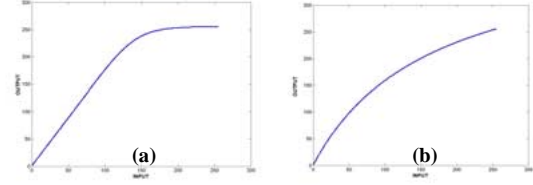


Figure 4. LUTs derived from curves with (a) A=7 and C=0.13; (b) A=0.85 and C=1.

2.3. Exposure Correction

Once the visually relevant regions (e.g. skin) are identified, the exposure correction is carried out using the mean gray value of those regions as reference point. A simulated camera response curve is used for this purpose, which gives an estimate of how light values falling on the sensor become final pixel values. Thus it is a function

$$f(q) = I \quad (7)$$

where q represents the ‘light’ quantity and I the final pixel value [1]. This function can be expressed ([4], [6]) by using a simple parametric closed form representation:

$$f(q) = \frac{255}{(1 + e^{-Aq})^C} \quad (8)$$

where parameters A, and C can be used to control the shape of the curve and q is supposed to be expressed in 2-based logarithmic unit (usually referred as “stops”). These parameters could be estimated, depending on the specific image acquisition device, using the techniques described in [7] or chosen experimentally. The offset from the ideal exposure is computed using the f curve and the average gray level of relevant regions avg , as:

$$\Delta = f^{-1}(128) - f^{-1}(avg) \quad (9)$$

The luminance value $Y(x,y)$ of a pixel (x,y) is modified as follows:

$$Y'(x,y) = f(f^{-1}(Y(x,y)) + \Delta) \quad (10)$$

Note that all pixels are corrected. Basically the previous step is implemented as a LUT transform (Fig.4 shows two correction curves with different A, C parameters). Final color reconstruction is done using the same approach

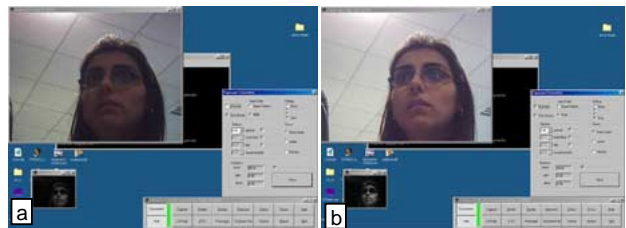


Figure 5. Framework interface for STV6500 - E01 EVK 502 VGA sensor: (a) before and (b) during real-time Skin Dependent Exposure Correction. The small window with black background represents the detected skin.

described in [10] to prevent relevant HUE shifts and/or color de-saturation:

$$R' = 0.5 \cdot \left(\frac{Y'}{Y} \cdot (R+Y) + R - Y \right) \quad (11)$$

$$G' = 0.5 \cdot \left(\frac{Y'}{Y} \cdot (G+Y) + G - Y \right) \quad (12)$$

$$B' = 0.5 \cdot \left(\frac{Y'}{Y} \cdot (B+Y) + B - Y \right) \quad (13)$$

where R, G, and B are the input color values.

3. EXPERIMENTS

The proposed solution has been tested using a large database of images acquired at various resolutions, with different acquisition devices, both in Bayer and RGB format. In the first case the algorithm was implemented in a real-time framework, using a CMOS-VGA sensor on the "STV6500 - E01" Evaluation Kit equipped with a "502 VGA sensor" [14]. Fig. 5 shows screen-shots of the working environment. In the RGB case the algorithm was implemented as a post-processing step. Examples showing input RGB data and corresponding processed images are shown in Fig.6.

4. CONCLUSIONS

An exposure correction algorithm for both RGB and Bayer pattern images has been presented. Experiments show the effectiveness of the techniques in both cases. The overall computational cost of the proposed solution is negligible, thus it is well suited for real time applications. Future work consists in the hardware implementation of the algorithm for a real-time environment. Future works will include a comparison between the results in the different color spaces.

5. REFERENCES

[1] S. Battiato, A. Castorina, M. Mancuso, "High Dynamic Range Imaging: Overview and Application", accepted for Publication on SPIE Journal of Electronic Imaging 2002;
 [2] B.E. Bayer, "Color Imaging Array", Patent 3,971,065-1976;
 [3] A. Bosco, S. Battiato, M. Mancuso, G. Spampinato, "Temporal Noise Reduction of Bayer Matrixed Data", In Proceedings of IEEE ICME'02, International Conference on Multimedia and Expo 2002, Lausanne, Switzerland, August 2002;
 [4] S.A. Bhukhanwala, T.V. Ramabadran, "Automated Global Enhancement Of Digitized Photographs", IEEE Transactions on Consumer Electronics, Vol. 40, No. 1, 1994;
 [5] M. Mancuso, S. Battiato, "An Introduction to the Digital Still Camera Technology", ST Journal of System Research, Vol.2, No.2, Dec.2001;
 [6] M. Mancuso, A. Bosco, S. Battiato, G. Spampinato, "Adaptive Temporal Filtering for CFA Video Sequences", In

Proceedings of IEEE ACIVS 2002, Advanced Concepts for Intelligent Vision Systems, pp. 19-24, Ghent University, Belgium, September 2002;
 [7] S. Mann, "Comparative Equations with Practical Applications in Quantigraphic Image Processing", IEEE Transactions on Image Processing, Vol. 9, No. 8, 2000;
 [8] G. Messina, S. Battiato, M. Mancuso, A. Buemi, "Improving Image Resolution by Adaptive Back-Projection Correction Techniques", IEEE Transaction on Consumer Electronics 2002, vol.48, no.3, pp.400-408, August 2002;
 [9] S.L. Phung, A. Bouzerdoum, D. Chai, "A Novel Skin Color Model In YCbCr Color Space And Its Application To Human Face Detection", International Conference on Image Processing 2002, vol.1, pp.289-292, September 2002;
 [10] S. Sakaue, A. Tamura, M. Nakayama, S. Maruno, "Adaptive Gamma Processing of the Video Cameras for the Expansion of the Dynamic Range", IEEE Transactions on Consumer Electronics, Vol. 41, No. 3, pp.555-561, August 1995;
 [11] M. Soriano, B. Martinkauppi, S. Huovinen and M. Laaksonen, "Skin Color Modeling Under Varying Illumination Conditions Using the Skin Locus for Selecting Training Pixels", Real-time Image Sequence Analysis - RISA2000, Aug.31-Sep.1, Finland, 2000;
 [12] J. Yang, W. Lu, and A. Waibel, "Skin-colour Modeling and Adaptation", Technical Report CMU-CS-97-146, School of Computer Science, Carnegie Mellon University, 1997;
 [13] B.D. Zarit, B.J. Super, and F.K.H. Quek, "Comparison of Five Colour Models in Skin Pixel Classification", Proc. of Int. Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, IEEE Computer Society, Corfu, Greece, pp. 58-63, 1999;
 [14] Colour Sensor Evaluation Kit VV6501, STMicroelectronics, Edinburgh, www.edb.st.com/products/image_sensors/501/6501evk.htm.



Figure 6. Experimental results: (a) original images acquired by Nokia 7650 VGA sensor compressed in Jpeg format; (b) Corrected output; (c) Image acquired with CCD Sensor (4.1 MPixel) Olympus E-10@ Camera; (d) Corrected output image.