Improving digital images by adaptive exposure

By Giuseppe Messina
Alfio Castorina
Angelo Bosco
Sebastiano Battiato
Imaging Group
STMicroelectronics Inc.

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

DSCs are currently among the devices most commonly employed for acquiring digital images. Sensors of greater resolutions having low-cost/low-consumption DSPs are readily available in commerce. This has led to the development of DSCs capable of acquiring images of considerable resolution and quality. Image improvement is obtained by increasing the resolution of the sensor and/or by using sophisticated image-processing algorithms (**Equation 1, Equation 3, Equation 6** and **Equation 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, where several factors contribute to acquire badly exposed pictures: poor optics, absence of flashgun, etc.

There is no exact definition of what a correct exposure should be. It is possible to abstract a generalization and define the best exposure that enables one to reproduce the most important regions (according to contextual or perceptive criteria) with a level of gray or brightness, more or less in the middle of the possible range. Our method is different from the one described in **Equation 4**, since the whole process can also be implemented directly on Bayer pattern images (**Equation 2**). Simpler statistical measures (focus and contrast) are used to identify information carrying regions.

Figure 1: 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.

Furthermore, a features-extraction method based on skin detection is introduced. The selection of image areas having particular characteristics allows application of selective exposure correction. These exposure adjustment techniques are designed essentially for mobile sensors applications.

Approach

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 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 are identified (for example, the pixels belonging to skin features), a global tone-correction technique is applied using the mean gray levels of identified regions as the main parameter.

Contrast and focus features: To identify regions of the image that contain more information, the luminance plane is subdivided into 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.

Focus characterizes the sharpness or edgeness of the block and is useful in identifying regions where HF components are present. If the said measures were simply computed on highly-underexposed images, then the regions having better exposure would always have higher contrast and edgeness compared to those that are obscured.

To perform a visual analysis revealing the most important features regardless to lighting conditions, a new "visibility' image" is constructed by pushing the mean gray 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. The contrast measure is computed by simply building a histogram for each block, and then calculating its deviation from the mean value. A high deviation value denotes good contrast. 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 gray-level $\tilde{I}[i]$.

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

Histogram deviation D is computed as:

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

where M is the mean value:

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

The focus measure is computed by convolving each block with a simple 3x3 Laplacian filter. To discard irrelevant HF pixels (mostly noise), the outputs of the convolution at each pixel are thresholded. The mean focus value of each block is computed as:

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

where N is the number of pixels and the thresh() operator discards values lower than a fixed threshold T. Once the values F and D are computed for all the blocks, relevant regions will be classified using a linear combination of both values.

Skin recognition: Most existing

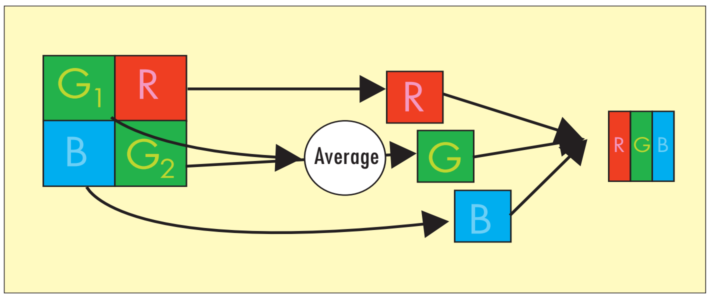

Figure 2: Bayer data sub-sampling generation.

methods for skin color detection usually threshold a skin-similarity measure for each pixel. Human skin colors are distinctive from the colors of the majority of other natural objects. It has been found that human skin colors are clustered in various color spaces. The skin color variations between people are mostly due to intensity differences. These variations can therefore be reduced using chrominance components only.

Equation 12 has demonstrated that the distribution of human skin colors can be represented by a 2D 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 –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 color class, given its chrominance vector ($\vec{x}$) is then represented by:

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

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

$$[d(\vec{x})]^2=(\vec{x}-\vec{\mu})\cdot\sum^{-1}(\vec{x}-\vec{\mu}) \qquad (6)$$

The value $d(\vec{x})$ determines the probability that a given block belongs to the skin color class. The larger the distance $d(\vec{x})$, the lower the probability that the block belongs to the skin color class. Due to the large quantity of color spaces, distance measures and 2D distributions, many skin recognition algorithms can be used. The proposed skin color algorithm is independent from the exposure correction. Thus, depending on the color space used (YCrCb or RGB), we introduce two alternative techniques aimed to recognize skin regions:

1. By using the input YCbCr image and the conditional probability (**Equation 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 (**Figure 1b**). The higher the gray 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).

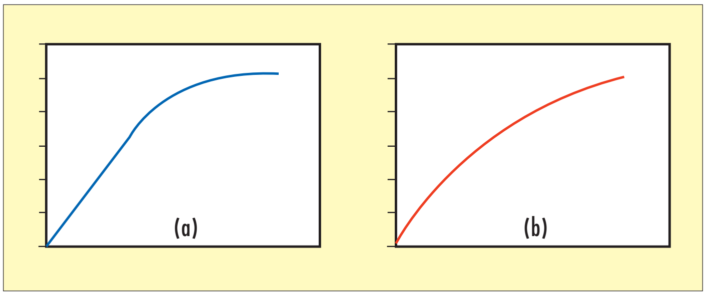

Figure 3: LUTs derived from curves with (a) A=7 and C=0.13; (b) A=0.85 and C=1.

Chrominance values representing skin are clustered in a specific area of the (r, g) plane, called skin locus, as defined in **Equation 11**. Pixels having a chrominance value belonging to the skin locus will be used to correct exposure (**Figure 1c**). 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 **Figure 2**. Each pixel of the quarter-sized RGB color image corresponds to a group of four pixels in the Bayer domain.

Once the visually relevant regions 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.

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

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

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

where parameters A and C can be used to control the shape of the curve, and q is supposed to be expressed in two-based logarithmic unit (usually referred as "stops"). These parameters could be estimated, depending on the specific image-acquisition device, using the techniques described in **Equation 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:

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

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

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

Note that all pixels are corrected. Basically, the previous step is implemented as an LUT transform (**Figure 3**) shows two correction curves with different A and C parameters). Final color reconstruction is done using the same approach described in **Equation 10** to prevent relevant HUE shifts and/or color desaturation:

$$R=0.5\left[\frac{Y'}{Y}(R+Y)+R-Y\right] \qquad (11)$$

$$G=0.5\left[\frac{Y'}{Y}(G+Y)+G-Y\right] \qquad (12)$$

$$B=0.5\left[\frac{Y'}{Y}(B+Y)+B-Y\right] \qquad (13)$$

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

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 Bayer 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. **Figure 4** shows screenshots of the working environment.

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. This will include a comparison between the results in the different color spaces. ☐

Figure 4: Framework interface for STV6500-E01 EVK 502 VGA sensor before and during real-time skin dependent exposure correction.