An Adaptive Global Enhancement Pipeline For Low Cost Imaging Sensors

Sebastiano Battiato, Alfio Castorina, Mirko Guarnera, Paolo Vivirito

Advanced System Technology Catania Lab - STMicroelectronics

E-mail :{Sebastiano.Battiato, Alfio.Castorina, Mirko.Guarnera, Paolo.Vivirito}@st.com

Abstract — *The paper describes a suitable algorithms pipeline able to enhance the picture quality of low cost imaging sensors, typically present in consumer devices (i.e. mobile phone, web-cams, PDA, etc) and accounts for compression artifacts, chromatic and geometric distortions, wrong exposure and poor sharpness.*

Index Terms — Adaptive sharpness, exposure correction, de-blocking, JPEG, white balancing.

I. INTRODUCTION

Typical consumer devices acquire digital images by properly using digital sensors (e.g. CCD/CMOS). Quality improvement is obtained by increasing the resolution of the sensor or by using more sophisticated image-processing algorithms, [1],[2] and [3]. For low cost imaging sensors the particular working conditions are not able to generate high quality images; several artifacts can be easily detected: compression artifacts due to the heavy compression, chromatic and geometric distortions and aberrations, color casting, etc. In this paper a suitable global post-enhancement pipeline, able to overcome such problems, is proposed. Experiments over a large data set of images acquired using real low-cost sensors have showed the real improvement obtained in term of perceived quality. The pipeline is shown in fig.1.
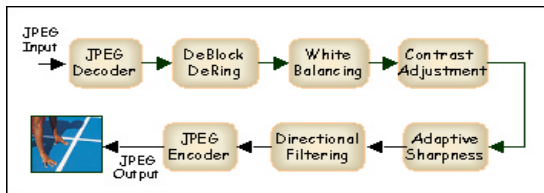

Figure 1 - The Proposed Pipeline.

II. THE OVERALL PIPELINE STEPS

A. De-blocking and De-ringing

Unpleasant JPEG artifacts are not unlikely to appear in compressed images; especially in rate-constrained transmission. Main artifacts are blocking and ringing. In order to attenuate both artifacts de-blocking and de-ringing filters were employed. Both filters were derived from the post-processing algorithms adopted in the MPEG4 [4] codec engine specifically adapted to work on JPEG compressed images.

Main differences regarded the de-blocking filter, where this meant operating on a block-by-block basis, and in guiding the adaptive filtering (DC offset mode for textured blocks and lowpass filter for smooth blocks) using the DC coefficient instead of the QP factor.

B. Geometric Distortion

Usually low cost cameras exhibits significant lens distortion, especially radial distortion. Geometric distortion describes the difference between the observed and real position of a point in the space. The radial symmetrical distortion can be corrected using a suitable polynomial relation [5]:

$$x' = x(1 + k_1 r^2 + k_2 r^4),$$
$$y' = y(1 + k_1 r^2 + k_2 r^4), \tag{1}$$

where $k_1 >> k_2$, (x,y) and (x',y') are the observed and ideal corrected image coordinates.

C. Automatic White Balancing

The role of white balancing is to compensate for color cast in scenes taken under non-white illumination. The algorithm here adopted can be seen as an improved variation from gray world algorithms, where the three channel energies are forced to be equal. Channel energies are computed by means of the following:

$$ce_c = \frac{\sum_{i=0}^{N-1} p_{c\_i} \cdot \min(p_{r\_i}, p_{g\_i}, p_{b\_i}) \cdot \boldsymbol{a}(T, p_{r\_i}, p_{g\_i}, p_{b\_i})}{\sum_{i=0}^{N-1} \boldsymbol{a}(T, p_{r\_i}, p_{g\_i}, p_{b\_i})} \quad c \in \{r, g, b\}, \tag{2}$$

where $p_{c\text{-}i}$ is the value of the c-channel of pixel i, N is the number of total pixels, and $\boldsymbol{a}$ is a 4-ary operator whose value is 0 if all channels are greater than an user fixed threshold T, 1 otherwise. Using the weighted average with respect to minimum values excessive saturation will be avoided, while the $\boldsymbol{a}$ discards useless near saturation values. Once the channel energies are computed they are normalized with respect to maximum energy providing 3 multiplicative corrective gains.

D. Exposure Correction

Acquiring scenes using consumer digital devices where no skilled settings are at hand, errors (i.e. wrong exposure) are not unlikely to occur. A post-processing step adapted from that described in [6] was used to correct such errors. Correction can be driven by the mean gray value of one of two kinds of visually relevant zones: a) zones where people are present,

thus suggesting a portrait; b) zones of high energy content, (i.e. focus, contrast) thus avoiding flat areas. If a relevant percentage of pixels are classified as skin the first operating mode is used (see [7] for more details), otherwise the second one. Skin detection has been done using the techniques described in [8], while zones of high-energy content are estimated using laplacian filters (e.g. focus) and histogram deviations (e.g. contrast). Corrections are done forcing the average gray value of the selected zones to occupy mid gray tonal range (e.g. near 130). Once offset Δ from desired exposure is computed, correction of all pixels Y channel is done using:

$$Y'(x,y) = f(f^{-1}(Y(x,y)) + \Delta) \qquad (3)$$

In the afore described formulas f is a kind of camera response function that could be parametrically described and estimated using the techniques in [9]. Y corrections are projected in the RGB space using the formulas described in [10].

E. Adaptive Sharpness

Low cost imaging sensors tend to exhibit, moving away from the center, an increasing degree of out-of-focus. In order to remove such blurring effect a simple linear combination between [11] and a classical un-sharp masking is used. Radial blurring effect is considered modulating the filtering proportionally to distance from the image center. Edge enhancement algorithms are typically affected by noise and ringing artifacts, which are reduced combining adaptively, restored image with the original one. More specifically, homogeneous areas (poorly interested by lens blur and sensitive to noise) can be selected from blurred image, while textured from restored one using the following relation:

$$F(x,y) = k * G(x,y) + (1-k) * R(x,y), 0 < k < 1 \quad (4)$$

where F is resulting image, G is the original image, R is the linear filtered image and k, used as a measure of texturization, is derived evaluating in a 8-neighborhood of the (x,y) pixel the differences between its value and its neighboring.

F. Directional Smoothing

Final images are adaptively filtered around edges to remove some residual JPEG distortions together with some artifacts coming from the above-described sharpening operators. Edges orientations are identified using the Sobel filters and each pixel is filtered around the prominent (max magnitude) direction in a 3×3 surround. Filtering is carried out by means of an elliptical shaped gaussian filter given by:

$$
\begin{aligned}
f(x,y,\mathbf{a}) &= he^{-\frac{\tilde{x}^2}{2\mathbf{s}_x^2} - \frac{\tilde{y}^2}{2\mathbf{s}_y^2}}, \\
\tilde{x} &= x\cos(\mathbf{a}) - y\sin(\mathbf{a}), \\
\tilde{y} &= x\sin(\mathbf{a}) + y\sin(\mathbf{a}),
\end{aligned}
\qquad (5)
$$

where $\mathbf{s}_x$, $\mathbf{s}_y$ are the variances along the two dimensions, h is a normalization constant and $\mathbf{a}$ is the orientation angle.

III. EXPERIMENTAL RESULTS

In order to evaluate the real performances of the proposed enhancement pipeline, a set of 300 images has been used. The input images acquired in different working conditions by a typical VGA sensor are compressed using the "low quality" JPEG setting of a 'typical' consumer mobile imaging phone. Experiments confirm the real perceived improvement in term of perceived quality (Color, Contrast/Sharpness, Artifact removal, etc.). Some examples are reported in Figure 2.

Figure 2 -Original (left) and processed (right) images.

IV. CONCLUSIONS AND FUTURE WORKS

An effective enhancement pipeline has been described. Future works will include the possibility to evaluate numerically the quality improvement obtained using some no-reference quality metric assessment.

REFERENCES

[1] A. Bosco, M. Mancuso, S. Battiato, G. Spampinato "*Temporal Noise Reduction of Bayer Matrixed Video Data*", Proceedings of IEEE ICME'02 International Conference on Multimedia and Expo 2002, pp.681-684, Lausanne, Switzerland, August 2002;

[2] 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;

[3] S. Battiato, A. Castorina, M. Mancuso "*High Dynamic Range Imaging: Overview and Application*", Accepted for publication: SPIE Journal of Electronic Imaging, November 2002;

[4] ISO/IEC JTC1/SC29/WG11 N 2502:Final Draft of International Standard MPEG-4;

[5] Z. Zhang. "*Flexible Camera Calibration By Viewing a Plane From Unknown Orientations*", International Conference on Computer Vision (ICCV'99), Corfu, Greece, pages 666-673, September 1999;

[6] S.A. Bhukhanwala, T.V. Ramabadran, "*Automated Global Enhancement Of Digitized Photographs*", IEEE Transactions on Consumer Electronics, Vol. 40, No. 1, 1994;

[7] S. Battiato, A. Bosco, A. Castorina, G. Messina, "*Automatic Global Image Enhancement by Skin Dependent Exposure Correction*", In Proceedings of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing – NSIP 2003 – Grado, Italy - June 2003;

[8] 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;

[9] S. Mann, "*Comparametric Equations with Practical Applications in Quantigraphic Image Processing*", IEEE Transactions on Image Processing, Vol. 9, No. 8, 2000;

[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] P. Vivirito, S. Battiato, S. Curti, M. La Cascia, R. Pirrone – "*Restoration of Out of Focus Images Based on Circle of Confusion Estimate*", Proceedings of SPIE 47th Annual Meeting 2002 – Applications of Digital Image Processing – Vol. 4790, Seattle, Washington, USA, July 2002.