2005

Novel color processing architecture for digital cameras with CMOS image sensors

Chaminda Weerasinghe  
Toshiba, Australia

Wanqing Li  
University of Wollongong, wanqing@uow.edu.au

Igor Kharitonenko  
University of Wollongong, igor@uow.edu.au

Magnus Nilsson  
Awapatent, Sweden

Sue Twelves  
Sensory Networks, Australia

Publication Details
Novel color processing architecture for digital cameras with CMOS image sensors

Abstract
This paper presents a color processing architecture for digital color cameras utilizing complementary metal oxide semiconductor (CMOS) image sensors. The proposed architecture gives due consideration to the peculiar aspects of CMOS image sensors and the human visual perception related to the particular application of digital color photography. A main difference between the proposed method and the conventional systems is the fact that color correction module is located before the interpolation module. Therefore, a method of performing color correction on a color filter array (CFA) pattern is also provided in this paper. The interpolation algorithm is especially designed to solve the problem of pixel cross talk among the pixels of different color channels. The algorithm separates the green channel into two planes, one highly correlated with the red channel and the other with the blue channel. These separate planes are used for red and blue channel interpolation. The implementation details related to managing four color channel values is also described. Experiments conducted on McBeth color chart and natural images have shown that the proposed color processing chain produces better quality images with improved SNR.

Disciplines
Physical Sciences and Mathematics

Publication Details

This journal article is available at Research Online: http://ro.uow.edu.au/infopapers/2832
Novel Color Processing Architecture for Digital Cameras with CMOS Image Sensors
Chaminda Weerasinghe, Wanqing Li, Igor Kharitonenko, Magnus Nilsson and Sue Twelves

Abstract — This paper presents a color processing architecture for digital color cameras utilizing Complementary Metal Oxide Semiconductor (CMOS) image sensors. The proposed architecture gives due consideration to the peculiar aspects of CMOS image sensors and the human visual perception related to the particular application of digital color photography. A main difference between the proposed method and the conventional systems is the fact that color correction module is located before the interpolation module. Therefore, a method of performing color correction on a color filter array (CFA) pattern is also provided in this paper. The interpolation algorithm is especially designed to solve the problem of pixel cross talk among the pixels of different color channels. The algorithm separates the green channel into two planes, one highly correlated with the red channel and the other with the blue channel. These separate planes are used for red and blue channel interpolation. The implementation details related to managing four color channel values is also described. Experiments conducted on McBeth color chart and natural images have shown that the proposed color processing chain produces better quality images with improved SNR.

Index Terms — color correction, color interpolation, color processing architecture, pixel cross talk.

I. INTRODUCTION

In single sensor electronic imaging systems, scene color is acquired by sub-sampling in three-color planes to capture color image data simultaneously for red, green and blue color components. Usually this is accomplished by placing a mosaic of red, green and blue filters over a 2D single sensor array. A method of arranging red green and blue pixels to form a mosaic pattern (e.g. Bayer pattern) [1] is shown in Fig. 1.

One significant characteristic of this type of sensors is horizontally adjacent pixels appear to contribute significantly to the response of their neighbors. Some colors exhibit a large variation between Gb and Gr values depending on the dominance of B and R channel values. This discrepancy gives rise to a blocking effect on the color interpolated image, and also spreading of false colors into detailed structures, due to R and B channel induced errors. There is a need to correct Gr/Gb discrepancy before further processing, in particular, due to the fact that the result of color interpolation may be adversely affected by the Gr/Gb discrepancy.

The existing technologies suggest modification of the captured green pixel values based on the adjacent pixel gradient classifications [2]. However, this method can seriously alter the captured scene information by way of excessive smoothing. It is also undesirable to discard the captured green pixel information, which cannot be recovered at a later stage in processing. The proposed method does not discard any of the captured green pixel information. Instead, it will create two green planes (i.e. Gr and Gb planes) by introducing additional green pixel values based on the neighborhood pixels. Due to the high correlation between R-Gr planes and B-Gb planes, the proposed method provides more accurate scene information (i.e. edges etc.) from Gr and Gb planes to R and B interpolated planes respectively.

Another method proposed comprise of applying a linear low pass filter on 5x5 pixel neighborhoods to eliminate the artifacts caused by the Gr/Gb difference [3]. However, this can excessively smooth the edges and thin elements will be removed from the scene. In order to sharpen the image, a matching high pass filter will have to be used. However, this can adversely affect the noise performance of the system.

II. PIXEL CROSS-TALK

There are several possible factors that may contribute to the cross talk. Optically, light may pass through one pixel filter at such an oblique angle that it strikes its adjacent pixels by the time it propagates down to the sensor surface. Electrically, sensor read-out circuits may allow for the signal read from one pixel to influence the signal read from another pixel. Architecturally, carriers generated by penetrating photons
under a pixel may diffuse to a nearby pixel depletion region and can be collected by the nearby pixel. The depth by which a photon will penetrate a silicon substrate before generating a carrier is strongly wavelength dependent [4] and the longer the wavelength, the deeper the penetration. As a result, the diffusion causes a strong cross talk between the red pixels and their Gr neighbors [5]. Color channel variation due to pixel cross talk is illustrated in Fig. 2. The $f_{Gr}(x)$ curve is usually above the $f_{Gb}(x)$ curve when $f_{R}(x) > f_{B}(x)$, which is the case for most colors, due to the high absorption of the blue channel through the CFA. The $f_{Gb}(x)$ curve is above the $f_{Gr}(x)$ in many shades of blue.

As can be seen in Figure 2, the rate of change in gradients of $f_{Gr}(x)$ and $f_{Gb}(x)$ correspond to that of $f_{R}(x)$ and $f_{B}(x)$ respectively. This phenomenon can be exploited to preserve details and reduce moiré artifacts along the sharp edges. Therefore, the relationship between the color-channels due to cross talk can be modeled as

\[ f_{R}^{\prime\prime}(x) = f_{Gr}^{\prime\prime}(x) \]

\[ f_{B}^{\prime\prime}(x) = f_{Gb}^{\prime\prime}(x) \]  \( (1) \)

\( (2) \)

Where $f^{\prime\prime}(x)$ represents the second order derivative.

Consider the following R – Gr line on the CFA,

\[ f_{Gr}(-2) \quad f_{R}(-1) \quad f_{Gr}(0) \quad f_{R}(+1) \quad f_{Gr}(+2) \]

The task at hand is the estimation of $f_{R}(0)$ under the condition that $f_{R}^{\prime\prime}(0) = f_{Gr}^{\prime\prime}(x)$ at $x = 0$. Using the second order Laplacian gradients, it is easy to prove that

\[ f_{R}(0) = \left[ \frac{f_{R}(-1) + f_{R}(+1)}{2} \right] + \left[ \frac{2f_{Gr}(0) - f_{Gr}(-2) - f_{Gr}(+2)}{4} \right] \]  \( (3) \)

This result can be used to interpolate $R$ and Gr channels in the horizontal direction. However, if interpolation needs to be performed in the vertical direction, this is impossible due to the fact that no Gr values reside in the same vertical line as R, and vice versa. Therefore it is necessary to estimate the $f_{Gr}(x)$ values in the positions of $f_{Gb}(x)$ and vice versa. It is also important to keep the Gr and Gb channels separated until the interpolation of R and B channels are completed to exploit the additional information provided by $f_{R}^{\prime\prime}(x) = f_{Gr}^{\prime\prime}(x)$ and $f_{B}^{\prime\prime}(x) = f_{Gb}^{\prime\prime}(x)$ relationships. When this is completed, the G channel can be estimated using $f_{Gr}(x)$ and $f_{Gb}(x)$ values, compensating for the cross talk. Pre-compensating for pixel cross talk in green channel as described in [5][6] eliminates the opportunity to exploit the inter-channel relationships.

To reconstruct $f_{G}(x)$ from $f_{Gr}(x)$ and $f_{Gb}(x)$, the error in gradient between $f_{G}(x)$ and the sampled $f_{Gr}(x)$ and $f_{Gb}(x)$ is minimized, in order to guarantee that the sharpness of the image is preserved [6]. Therefore, the optimization criterion can be expressed as shown in Equation (4)

\[ f_{G}(x) = \min \left[ \int \left| \nabla f_{Gr}(x) - \nabla f_{Gb}(x) \right|^{2} \right] dx \]  \( (4) \)

To solve Equation (4) further assumptions are required about the G curve. Reasonable assumptions include the local average of the G curve is close to either the Gr or Gb curve or is between the Gr and Gb curves.

In the former case, Equation (4) can be solved subject to

\[ \tilde{f}_{G}(x) = \tilde{f}_{Gr}(x) \]

\[ \tilde{f}_{G}(x) = \tilde{f}_{Gb}(x) \]  \( (5) \)

\( (6) \)

Where $\tilde{f}_{G}(x)$, $\tilde{f}_{Gr}(x)$, and $\tilde{f}_{Gb}(x)$ are local averages around $x$. In the latter case, Equation (4) can be solved subject to

\[ \tilde{f}_{G}(x) = \left( \tilde{f}_{Gr}(x) + \tilde{f}_{Gb}(x) \right) / 2 \]

\[ (7) \]

The local average can be estimated from the neighborhood of a pixel centered at $x$.

III. PROPOSED COLOR PROCESSING CHAIN ARCHITECTURE

The proposed color processing chain is shown in Fig. 3. It consists of the following modules:

1) **Pre-processing module**: The green channel is split into two planes, which separates the Gr and the Gb values. All subsequent processing is carried out separately for the red channel and the blue channel using the associated green channel values from the Gr plane and GB plane respectively.

2) **Color correction module**: The color correction is performed on the Bayer pattern using a neighborhood-based algorithm.

3) **Color Interpolation Module**: The interpolation algorithm takes into account the correlation among the different color channels.

The advantage of this system is that it provides a better quality output with minimum noise escalation using the proposed architecture. However, the splitting of the green channel into separate Gr and Gb channels create non-standard interfaces between the various modules.
A. Pre-processing Module

The proposed method employs a simple median filter among the 4 Gr values and the single Gb value (see Figure 4) to replace the existing Gb value. This creates a GR plane which favours the R channel. However, B channel favored image details can be lost in this process. It is important to preserve the captured pixel values, and also avoid unnecessary smoothing of image structures. Therefore, the process needs to be repeated with the roles of Gr and Gb swapped, creating separate GR and GB planes. It should be noted that in this process, the captured green pixel values will always be preserved in either GR plane or in the GB plane [7].

![Bayer pattern](image)

**Fig. 3. Proposed Color Processing Chain Architecture.**

B. Color Correction Module

Spectral sensitivity of solid-state color image sensors usually differs from ideal color matching functions. To improve the color reproduction accuracy, digital cameras perform color correction. A widely used solution is to use a 3x3 color correction matrix with the coefficients optimized on a number of color samples. Although this method achieves a significant improvement in color rendering, it also amplifies noise. Therefore, the noise level of the color-corrected images becomes substantially higher. The proposed color correction algorithm reduces amplification of the noise. As CMOS image sensors are inherently noisy, this algorithm is especially beneficial for use with such sensors [8]. The color correction is often carried out using the following matrix:

\[
\begin{bmatrix}
R_c \\
G_c \\
B_c
\end{bmatrix} =
\begin{bmatrix}
a_{11} & -a_{12} & -a_{13} \\
-a_{21} & a_{22} & -a_{23} \\
-a_{31} & -a_{32} & a_{33}
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

(8)

As shown in Equation (8), subtracting weighted values of the other two channels performs the color correction for a particular channel. For example, red channel can be corrected using Equation (9).

\[
R_c = a_{11}R - a_{12}G - a_{13}B
\]

(9)

However, in the proposed color correction algorithm formulated in Equation (10), instead of using the values of the current pixels to make the adjustments, the average of the various color planes in the neighborhood is used.

\[
\begin{bmatrix}
R_c \\
G_c \\
B_c
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
+ \begin{bmatrix}
a'_{11} & -a'_{12} & -a'_{13} \\
-a'_{21} & a'_{22} & -a'_{23} \\
-a'_{31} & -a'_{32} & a'_{33}
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

(10)

Where \(a'_{11} = a_{11} - 1\), \(a'_{22} = a_{22} - 1\) and \(a'_{33} = a_{33} - 1\). However, this calculation does not need to be performed directly, for example, \(R_c\) can be calculated as shown in Equation (11)

\[
R_c = R + a_{11}R - a_{12}G - a_{13}B
\]

(11)

Since the G channel is already separated in to Gr and Gb channels, \(\bar{G}\) is represented by \(\bar{G}_R\) when computing \(R_c\), and by \(\bar{G}_B\) when computing \(B_c\).

This color correction method uses three low-pass filters (LPF) via averaging, to prevent noise amplification. It is well known that a straightforward application of the LPF has a negative impact on the image sharpness. This method avoids this undesirable side effect of noise reduction by applying the LPF in a non-conventional way. As a result, the blur introduced by the LPF is naturally compensated through combining data from several color channels. Thus, whilst suppressing the noise, this method of color correction also preserves the original sharpness of the images. Operating directly on the Bayer pattern, the system processes three times fewer samples than that of the interpolated image.

Since the average values can be calculated directly from the Bayer pattern, color correction can be carried out prior to interpolation. Fig. 5 shows the channel information available for a 5x5 window on a Bayer pattern image. As can be seen, in this particular window, there are 12 green, 9 red and 4 blue values available for average calculations.

In this particular example, where an R-value is being filtered, the G values would be obtained from the Gr plane produced during the pre-processing stage. When a B value is being filtered, the G values are taken from the Gb plane.

![Median filter](image)

**Fig. 4. Separation of green channel into the GR and GB planes.**

![Average data from a 5x5 window on a Bayer pattern](image)

**Fig. 5. Average data from a 5x5 window on a Bayer pattern.**
C. Interpolation Module

There are two stages to the interpolation. Firstly, the missing green pixel values for the GR and GB planes are interpolated separately using a suitable interpolation algorithm, such as bilinear interpolation with second order Laplacian correction terms [9][10]. Secondly, the R and B pixels are interpolated. High correlation between R-GR and B-GB planes can be exploited to interpolate R and B channels to gain additional details. Therefore, R is interpolated using the edge information of GR, and, B using GB. The color interpolated image is composed using the R, B and average values of GR and GB planes [7]. Complete block diagram of the proposed interpolation scheme is shown in Fig. 6.

![Block diagram of the proposed color interpolation method.](image)

Fig. 6. Block diagram of the proposed color interpolation method.

Four green pixels are used to decide whether to interpolate in the horizontal direction or the vertical direction. It is undesirable to interpolate across an edge because this would have an effect of smoothing the edges and losing image sharpness. Therefore, the gradient is checked between two vertical green pixels and two horizontal green pixels and the direction of the lowest gradient is selected. R interpolation is performed using Equation (3). Similar equations are used for interpolating B plane. Consequent to performing interpolation on R and B, the G plane values are computed using the available GR and GB values using the Equation (7).

IV. IMPLEMENTATION

Implementation of the algorithms described in section III can involve three key efficiency measures:

1. Memory management
2. Generic implementation that deals with all four different types of Bayer patterns
3. Reuse of mean calculations on a GRBG block for color correction.

A. Memory Management

The separation of the green channel into two separate GR and GB planes requires a total of 5 times the image size to store the information: $1 \times R$, $1 \times B$, $3 \times G$ planes. However, this extra memory requirement can be reduced, by using the fact that both the GR and the GB planes only occupy half the space initially until the interpolation phase.

Key for Fig. 7 is as follows:

- Gr - green pixels on a Bayer pattern red row
- Gb - green pixels on a Bayer pattern ‘blue’ row
- Grec - estimated Gr pixels for ‘GR’ plane
- Gbec - estimated Gb pixels for ‘GB’ plane
- Grc - original Gr pixels, which have been color corrected
- Gbe - original Gb pixels, which have been color corrected
- Gbec - estimated Gr pixels, which have been color corrected
- Gb - original red pixels
- Rc - color corrected red pixels
- Bc - color corrected blue pixels
- G - original green pixels
- R - original red pixels
- B - original blue pixels
Fig. 7 shows how the GR and GB plane information is stored in the red and blue storage areas respectively by using their empty spaces during the pre-processing stage. When the color correction stage takes place, the color corrected GR pixels are stored in the green area whereas the color corrected GB pixels are moved from the blue area to the red area. The reason for having to swap the GB pixels over from the blue to red area is because the process of color correction requires the pre-processed pixels prior to color correction. Therefore, it is not possible to overwrite the values whilst the correction is taking place.

During the interpolation stage, a new GB plane is used to store the interpolated GB data. The original green area is used to store the interpolated GR data. The red and blue areas are then free to store the interpolated values for the red and blue pixels.

The advantage of this memory management scheme is that only one extra image plane is required. In addition to this, the extra plane is embedded within the interpolation algorithm and is not necessary for any other stage.

C. Reuse of the statistical means on a GRBG block for color correction

The color correction algorithm can be implemented more efficiently by treating a GR/GB Bayer pattern block as a single entity, see Fig. 9.

In this efficiency measure, the color correction algorithm processes a GR/GB block altogether. This is accomplished by using the same mean values to adjust each of the 4 color channels. The mean values are calculated from the same 4x4 matrix as shown in Fig. 9. By only having to calculate the mean values once for each block, the color correction for four channel values can be carried out in one pass, thus allowing much faster processing speeds.

V. EXPERIMENTAL RESULTS

The following results are included to illustrate the capability of the proposed method in terms of, false color suppression, blocking artifact reduction and visible noise suppression, compared to the interpolation methods currently available [10]. It should be noted that results are produced for the specific case of images captured via a CMOS sensor with Bayer Color filter array (CFA). No gamma correction has been applied to the processed images to illustrate the performance of the proposed modules only. The conventional color processing chain used for performance comparison is shown in Figure 10.

In order to compare the two different architectures shown in Figure 3 and Figure 10, same interpolation equations (i.e. similar to Equation (3)), and same color correction matrix coefficients are used. However, in the conventional architecture, no Gr/Gb separation is performed.
A visual comparison between images processed using conventional and proposed color processing methods is performed. Zoomed (x 2) relevant sections of the images are shown in Fig. 11, whereas the full images (640 x 480 pixels) are given in Fig. 12.

It was observed that, on average, there is 54% reduction of false color values specially related to the green channel. This is a significant improvement, since the human eyes are more sensitive to the green channel. The processed images of the McBeth Color Chart (shown in Figure 12) were used to derive the above conclusion.

A Quantitative comparison on signal to noise ratio (SNR) of output images was performed using the McBeth Color Chart. The results are shown in Table 1.

**Table I**

<table>
<thead>
<tr>
<th>Color</th>
<th>Noise Suppression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>0.7% - 3.78%</td>
</tr>
<tr>
<td>Green</td>
<td>42.1% - 64.8%</td>
</tr>
<tr>
<td>Blue</td>
<td>1.75% - 3.17%</td>
</tr>
</tbody>
</table>

Therefore, the proposed interpolation method suppresses false colors both in smooth regions and in thin edge regions, reduces blocking artifacts caused by Gr/Gb difference and, significantly suppresses visible noise in the processed image.
VI. Conclusion

A color processing chain in a single sensor color camera using green channel separation was presented. Green channel separation can be implemented with any interpolation scheme utilizing edge sensing, pattern recognition or green plane correction terms. The results indicate that the proposed method minimizes blocking artifacts, false colors and visible noise, producing general improvement on the image quality.

REFERENCES


Chaminda Weerasinghe received BE Honors Class 1 with university medal from University of Wollongong, Australia in 1994 and his Ph.D. in image processing from University of Sydney, Australia in 1999. He is a recipient of many academic awards and medals from IEE, IEAust and IESA. Dr. Weerasinghe was with Motorola Australian Research Center (2000 – 2003) as a senior research engineer. He is currently with Toshiba (Australia) Pty. Ltd. (R&D Division). His main research interests are in color image processing; CMOS image sensors, surveillance camera systems, stereoscopic/panoramic video generation/display and raster image processing for printers.

Wanqing Li received B.Sc. in physics and electronics and M.Sc. in computer science from Zhejiang University, China in 1983 and 1987 respectively. In 1997, he received PhD in electronic engineering from The University of Western Australia, Australia. He was a lecturer from 1987 to 1990 and associate professor from 1991 to 1992, both with department of computer science and technology, Zhejiang University of China. From 1992 to 1993, he was a visiting research fellow with computer science department, Murdoch University, Australia. From 1997 to 1998, he worked as an Information Technology Officer with Bureau of Meteorology, Australia. He was with Motorola Australian Research Centre from 1998 to 2003 as a principal research engineer. He is currently with University of Wollongong. His research interests include computer vision, image processing and analysis, pattern recognition and neural networks.

Igor Kharitonenko was born in Odessa, Ukraine. He received the B.S. with honors in electronics engineering and Ph.D. degree from Odessa Polytechnic University in 1985 and 1993 respectively. Dr. Kharitonenko was a principal research engineer at Motorola Australian Research Centre (1996 – 2003) working on technology development for digital cameras and mobile video communicators. He is currently with University of Wollongong (Australia). His research interests include machine vision, CMOS image sensor architectures, image and video compression.

Magnus Nilsson was born in Gothenburg, Sweden in 1976. He received his BSc (electrical engineering) in 2000 from Chalmers University of Technology, Sweden, and his Master of Computer and Information Engineering in 2002 from Griffith University, Australia. He was working as a senior research engineer at Motorola Australian Research Center in Sydney, Australia (2001 – 2003). His research interests include Smart CMOS image sensors and embedded speech and image processing.

Sue Twelves received her B.Sc. Honors in physics from Southampton University in 1981, her M.Sc. in communications engineering from Imperial College in 1988 and her Ph.D. degree from James Cook University of North Queensland in 1998. Dr. Twelves was a senior research engineer at Motorola Australian Research Center, working on technology development for digital cameras, image compression and DSP applications. Sue Twelves is currently with Sensory Networks, Australia. Her research interests include signal and image processing for the telecommunications industry.