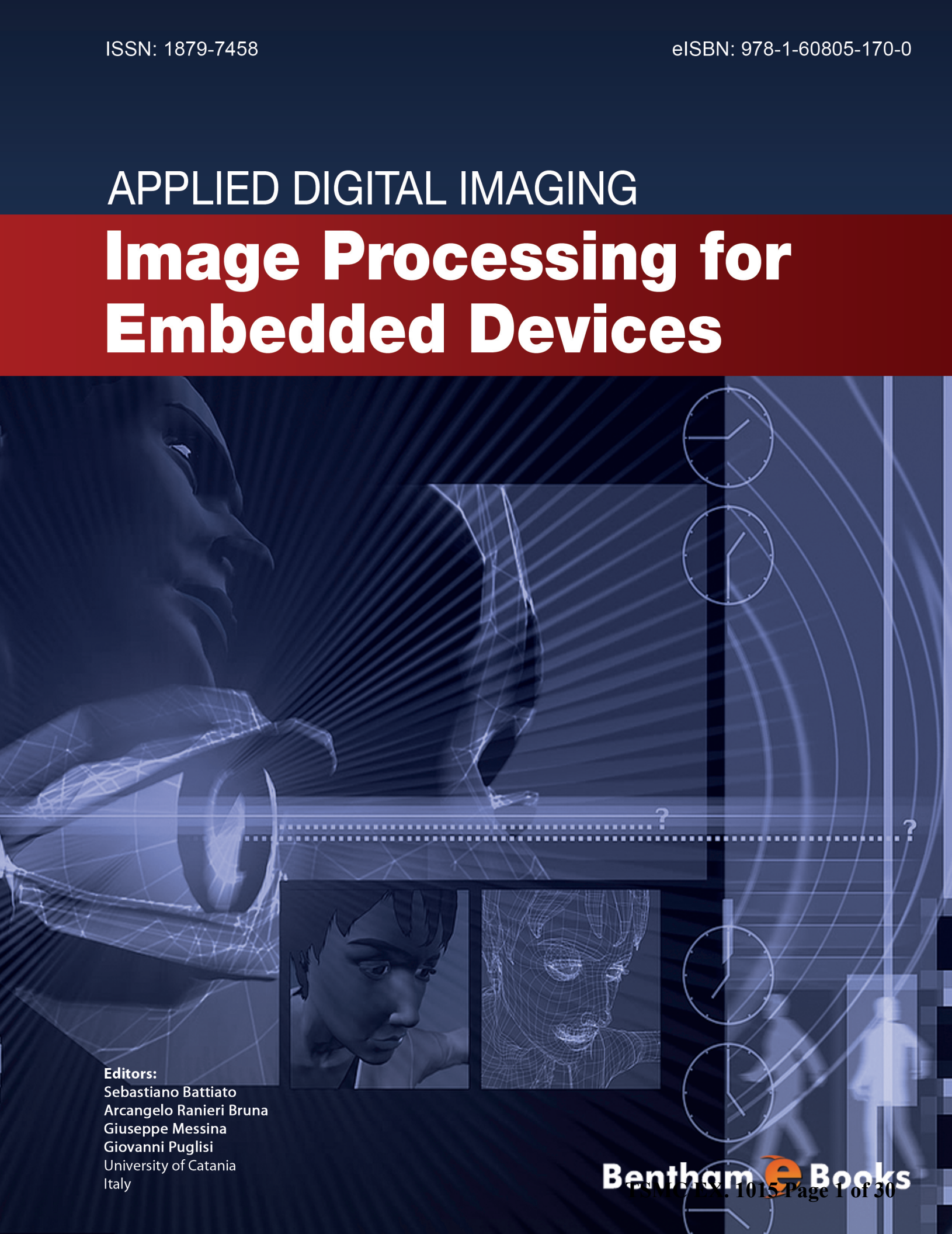


APPLIED DIGITAL IMAGING

Image Processing for Embedded Devices



Editors:
Sebastiano Battiato
Arcangelo Ranieri Bruna
Giuseppe Messina
Giovanni Puglisi
University of Catania
Italy

Image Processing for Embedded Devices

From CFA data to image/video coding

Editors:

Sebastiano Battiato
Arcangelo Ranieri Bruna
Giuseppe Messina
Giovanni Puglisi

eBooks End User License Agreement

Please read this license agreement carefully before using this eBook. Your use of this eBook/chapter constitutes your agreement to the terms and conditions set forth in this License Agreement. Bentham Science Publishers agrees to grant the user of this eBook/chapter, a non-exclusive, nontransferable license to download and use this eBook/chapter under the following terms and conditions:

1. This eBook/chapter may be downloaded and used by one user on one computer. The user may make one back-up copy of this publication to avoid losing it. The user may not give copies of this publication to others, or make it available for others to copy or download. For a multi-user license contact permission@bentham.org
2. All rights reserved: All content in this publication is copyrighted and Bentham Science Publishers own the copyright. You may not copy, reproduce, modify, remove, delete, augment, add to, publish, transmit, sell, resell, create derivative works from, or in any way exploit any of this publication's content, in any form by any means, in whole or in part, without the prior written permission from Bentham Science Publishers.
3. The user may print one or more copies/pages of this eBook/chapter for their personal use. The user may not print pages from this eBook/chapter or the entire printed eBook/chapter for general distribution, for promotion, for creating new works, or for resale. Specific permission must be obtained from the publisher for such requirements. Requests must be sent to the permissions department at E-mail: permission@bentham.org
4. The unauthorized use or distribution of copyrighted or other proprietary content is illegal and could subject the purchaser to substantial money damages. The purchaser will be liable for any damage resulting from misuse of this publication or any violation of this License Agreement, including any infringement of copyrights or proprietary rights.

Warranty Disclaimer: The publisher does not guarantee that the information in this publication is error-free, or warrants that it will meet the users' requirements or that the operation of the publication will be uninterrupted or error-free. This publication is provided "as is" without warranty of any kind, either express or implied or statutory, including, without limitation, implied warranties of merchantability and fitness for a particular purpose. The entire risk as to the results and performance of this publication is assumed by the user. In no event will the publisher be liable for any damages, including, without limitation, incidental and consequential damages and damages for lost data or profits arising out of the use or inability to use the publication. The entire liability of the publisher shall be limited to the amount actually paid by the user for the eBook or eBook license agreement.

Limitation of Liability: Under no circumstances shall Bentham Science Publishers, its staff, editors and authors, be liable for any special or consequential damages that result from the use of, or the inability to use, the materials in this site.

eBook Product Disclaimer: No responsibility is assumed by Bentham Science Publishers, its staff or members of the editorial board for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products instruction, advertisements or ideas contained in the publication purchased or read by the user(s). Any dispute will be governed exclusively by the laws of the U.A.E. and will be settled exclusively by the competent Court at the city of Dubai, U.A.E.

You (the user) acknowledge that you have read this Agreement, and agree to be bound by its terms and conditions.

Permission for Use of Material and Reproduction

Photocopying Information for Users Outside the USA: Bentham Science Publishers Ltd. grants authorization for individuals to photocopy copyright material for private research use, on the sole basis that requests for such use are referred directly to the requestor's local Reproduction Rights Organization (RRO). The copyright fee is US \$25.00 per copy per article exclusive of any charge or fee levied. In order to contact your local RRO, please contact the International Federation of Reproduction Rights Organisations (IFRRO), Rue du Prince Royal 87, B-1050 Brussels, Belgium; Tel: +32 2 551 08 99; Fax: +32 2 551 08 95; E-mail: secretariat@ifrro.org; url: www.ifrro.org This authorization does not extend to any other kind of copying by any means, in any form, and for any purpose other than private research use.

Photocopying Information for Users in the USA: Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Bentham Science Publishers Ltd. for libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Services, provided that the appropriate fee of US \$25.00 per copy per chapter is paid directly to Copyright Clearance Center, 222 Rosewood Drive, Danvers MA 01923, USA. Refer also to www.copyright.com

CONTENT

<i>Foreword</i>	<i>i</i>
<i>Preface</i>	<i>ii</i>
<i>Biographies</i>	<i>iv</i>
<i>Contributors</i>	<i>vii</i>
<i>Acknowledgement</i>	<i>viii</i>
CHAPTER	
1. Fundamentals and HW/SW Partitioning	01
<i>S. Battiato, G. Puglisi, A. Bruna, A. Capra and M. Guarnera</i>	
2. Notions about Optics and Sensors	10
<i>A. Bruna, A. Capra, M. Guarnera and G. Messina</i>	
3. Exposure Correction	34
<i>A. Castorina and G. Messina</i>	
4. Pre-acquisition: Auto-focus	54
<i>A. Capra and S. Curti</i>	
5. Color Rendition	92
<i>A. Bruna and F. Naccari</i>	
6. Noise Reduction	117
<i>A. Bosco and R. Rizzo</i>	
7. Demosaicing and Aliasing Correction	149
<i>M. Guarnera, G. Messina and V. Tomaselli</i>	
8. Red Eyes Removal	191
<i>G. Messina and T. Meccio</i>	
9. Video Stabilization	217
<i>T. Meccio, G. Puglisi and G. Spampinato</i>	
10. Image Categorization	237
<i>G. M. Farinella and D. Ravi</i>	
11. Image and Video Coding and Formatting	270
<i>A. Bruna, A. Buemi, G. Spampinato</i>	
12. Quality Metrics	310
<i>Ivana Guarneri</i>	
13. Beyond Embedded Devices	343
<i>S. Battiato, A. Castorina and G. Messina</i>	
Index	374

Foreword

Image Processing in embedded devices has been an area of growing interest with the revolution of digital imaging devices since the last decade of the 20th century and it will continue to expand to new frontiers in this century. Despite its relevance, there is not, as far as I know, a comprehensive publication that address this topic encompassing practical aspects of image processing design.

With chapters contributed by both experienced researchers from academia as well as researchers and engineers from industry, the present publication covers fundamental aspects of image processing in embedded devices such as exposure correction, auto-focus, color rendition, noise reduction, demosaicing, encoding, red-eye removal, image categorization and presents relevant quality metrics and also recent trends in imaging.

The editors have done an excellent job of bringing out contributors that work with the challenges of finding solutions and also implementing image processing solutions for embedded imaging devices in a daily basis with continuous spread across all relevant operational aspects for an imaging system.

I believe, the present publication is going to be beneficial not only to imaging and engineering students but also be a reference for academic researchers and engineers working in imaging industry.

This publication is also unique because it moves away from the traditional paper book for technical publications and follows the trend of electronic book. This makes the publication more accessible, more portable with current e-readers in the market, potentially more environment friendly without ever going out of print. Electronic publications have also attribute such language accessibility by electronic translations and text-to-speech software capabilities.

It is a great pleasure for me to write a foreword for this prestigious, multi-authored, international publication on a topic that I believe is very relevant to the imaging industry. Finally, I would like to compliment the editors and contributors for their effort in making this publication a great success.

Francisco Imai, Ph.D.

Principal Scientist

Canon Development Americas, Inc.

3300 North First Street

San Jose, CA, 95134 USA

Preface

Embedded imaging devices, such as digital still and video cameras, mobile phones, personal digital assistants, and visual sensors for surveillance and automotive applications, make use of the single-sensor technology approach. An electronic sensor (Charge Coupled Device - CCD or Complementary Metal-Oxide-Semiconductor - CMOS) is used to acquire the spatial variations in light intensity and then uses image processing algorithms to reconstruct a color picture from the data provided by the sensor. Acquisition of color images requires the presence of different sensors for different color channels. Manufacturers reduce the cost and complexity by placing a color filter array (CFA) on top of a single sensor, which is basically a monochromatic device, to acquire color information of the true visual scene.

The overall performance of any device are the result of a mixture of different components including hardware and software capabilities and, not ultimately, overall design (i.e., shape, weight, style, etc.).

This book is devoted to cover algorithms and methods for the processing of digital images acquired by single-sensor imaging devices. Typical imaging pipelines implemented in single-sensor cameras are usually designed to find a trade-off between sub-optimal solutions (devoted to solve imaging acquisition) and technological problems (e.g., color balancing, thermal noise, etc.) in a context of limited hardware resources. State of the art techniques to process multichannel pictures, obtained through color interpolation from CFA are very advanced. On the other hand, not too much is known and published about the application of image processing techniques directly on CFA images, i.e. before the color interpolation phase.

The various chapters of the book cover all aspects of algorithms and methods for the processing of digital images acquired by imaging consumer devices. More specifically, we will introduce the fundamental basis of specific processing into CFA domain (demosaicing, enhancement, denoising, compression). Also ad-hoc matrixing and color balancing techniques devoted to preprocess input data coming from the sensor will be treated. In almost all cases various arguments have been presented in a tutorial way in order to provide to the readers a comprehensive overview of the main basis of each involved topics. All contributors are well renowned experts in the field as demonstrated by the number of related patents and scientific publications.

The main part of the book analyzes the various aspects of the imaging pipeline from the CFA data to image and video coding. A typical imaging pipeline is composed by two functional modules (pre-acquisition and post-acquisition) where the data coming from the sensor in the CFA format are properly processed. The term pre-acquisition is referred to the stage in which the current input data coming from the sensor are analyzed just to collect statistics useful to set parameters for correct acquisition.

The book also contains a number of chapters that provide solution and methods to address some undesired drawbacks of acquired images (e.g., red-eye, jerkiness, etc.); an overview of the current technologies to measure the quality of an image is also given. Just considering the impressive (and fast) growth in terms of innovation and available technology we conclude the book just presenting some example of solution that makes

use of machine learning for image categorization and a brief overview of recent trends and evolution in the field.

Catania (Italy), June 2010.

Sebastiano Battiato
Arcangelo Ranieri Bruna
Giuseppe Messina
Giovanni Puglisi

Biographies

Sebastiano Battiato

Sebastiano Battiato was born in Catania, Italy, in 1972. He received the degree in Computer Science (summa cum laude) in 1995 and his Ph.D. in Computer Science and Applied Mathematics in 1999. From 1999 to 2003 he has led the "Imaging" team c/o STMicroelectronics in Catania. Since 2004 he has been a researcher at Department of Mathematics and Computer Science of the University of Catania. His research interests include image enhancement and processing, image coding and camera imaging technology. He published more than 90 papers in international journals, conference proceedings and book chapters. He has authored 2 books and is a co-inventor of about 15 international patents. He is a reviewer for several international journals and he has been regularly a member of numerous international conference committees. He has participated in many international and national research projects. He is an Associate Editor of the SPIE Journal of Electronic Imaging (Specialty: digital photography and image compression). He is a director (and cofounder) of the International Computer Vision Summer School. He is a Senior Member of the IEEE. For more details see (<http://www.dmi.unict.it/battiato>)

Arcangelo R. Bruna

Arcangelo R. Bruna received the degree in Electronic Engineering (summa cum laude) in 1998 at the University of Palermo. First he worked in a telecommunication company in Rome. He joined STMicroelectronics in 1999 where he works in the Advanced System Technology (AST) Catania Lab - Italy. Today he leads the Image Generation Pipeline and Codecs group and his research interests are in the field of image acquisition, processing and enhancement. He published several patents and papers in international conferences and journals.

Giuseppe Messina

Giuseppe Messina was born in Crange, France, in 1972. He received his MS degree in Computer Science in 2000 at the University of Catania doing a thesis about Statistical Methods for Textures Discrimination. Since March 2001 he has been working at STMicroelectronics in the Advanced System Technology (AST) Imaging Group as Software Design Senior Engineer II / PL. Since 2007 he is Ph.D. student in Computer Science at the University of Catania accomplishing a research in Information Forensic by Image/Video Analysis. He is member of the Image Processing Laboratory, at the University of Catania. His research interests are in the field of Image Analysis e Image Quality Enhancement. He is author of about several papers and patents in Image Processing field. He is a reviewer for several international journals and international conferences. He is an IEEE member.

Giovanni Puglisi

Giovanni Puglisi was born in Acireale, Italy, in 1980. He received his degree in Computer Science Engineering (summa cum laude) from Catania University in 2005 and his Ph.D. in Computer Science in 2009. He is currently contract researcher at the Department of Mathematics and Computer Science and member of IPLab (Image Processing Laboratory) at the University of Catania. His research interests include video stabilization, artificial mosaic generation, animal behavior and raster-to-vector conversion techniques. He is the author of several papers on these activities.

Image Processing Lab (<http://iplab.dmi.unict.it>)

IPLab research group is located at Dipartimento di Matematica ed Informatica in Catania. The scientific knowledge of the group is on Computer Graphics, Multimedia, Image processing, Pattern Recognition and Computer Vision. The group has a good expertise in the overall digital camera pipeline (e.g., acquisition and post acquisition processing) as well as a good and in-depth knowledge of the recognition of scene categorization field. This is confirmed by the numerous research paper, within the area of image processing in single sensor domain (in acquisition and post acquisition time) as well as different works relatively the semantic analysis of images content, to drive some image processing tasks such as image enhancement. Moreover, the collaboration between members of the Catania unit and industrial company leaders in single sensor imaging (e.g., STMicroelectronics) has already done the possibility of transferring to the industry (pre-competitive research) the knowledge acquired in academic research facilitating the industry in producing new advanced products and patents. A joint research lab IPLab-STMicroelectronics, has been recently created where researchers coming from both partners work together on imaging research topics. More specifically, 2 Ph.D. students in Computer Science (XXIII Ciclo Dottorato in Informatica - Università di Catania) have received financial support by STMicroelectronics to investigate about "Methodologies and Algorithms for Image Quality Enhancement for Embedded Systems". The group published more than 100 papers on topics related to the previous mentioned disciplines. Moreover the IPLab group established a number of international relationships with academic/industrial partners for research purpose. In the last years the group organized the "Fourth Conference Eurographics Italian Chapter 2006" and the "International Computer Vision Summer School 2007, 2008, 2009, 2010" (<http://www.dmi.unict.it/icvss>).

Advanced System Technology - Catania Lab - STMicroelectronics (<http://www.st.com>)

Advanced System Technology (AST) is the STMicroelectronics organization in charge of system level research and innovation. Active since 1998, AST responds to the need to strengthen the position of STMicroelectronics as a leading-edge system on chip company. The AST Catania Lab and, in particular, the Imaging Group, works on research and innovation in the field of imaging processing. Its mission is to acquire digital pictures with superior Performance/Cost using advanced image processing methodologies, to extend the acquisition capability of imaging devices through the development of new applications and to determine the computational power, the required bandwidth, the flexibility and the whole imaging engine. Its members have long experience in image algorithms, documented also by many patents and scientific publications. Primarily, through active contacts and collaborations with several universities and a dedicated joint lab with the IPLab of Catania University, they have concretized and made effective the link between academic and industrial R&D.

List of Contributors

Image Processing Lab, University of Catania, Italy.

Sebastiano Battiato	battiato@dmf.unict.it
Giovanni Maria Farinella	gfarinella@dmf.unict.it
Tony Meccio	meccio@dmf.unict.it
Giovanni Puglisi	puglisi@dmf.unict.it
Rosetta Rizzo	rosetta.rizzo@dmf.unict.it

Advanced System Technology - Catania Lab, STMicroelectronics, Italy.

Angelo Bosco	angelo.bosco@st.com
Arcangelo Ranieri Bruna	arcangelo.bruna@st.com
Antonio Buemi	antonio.buemi@st.com
Alessandro Capra	alessandro.capra@st.com
Alfio Castorina	alfio.castorina@st.com
Salvatore Curti	salvatore.curti@st.com
Mirko Guarnera	mirko.guarnera@st.com
Ivana Guarneri	ivana.guarneri@st.com
Giuseppe Messina	giuseppe.messina@st.com
Filippo Naccari	filippo.naccari@st.com
Daniele Ravi	daniele.ravi-uni@st.com
Giuseppe Spampinato	giuseppe.spampinato@st.com
Valeria Tomaselli	valeria.tomaselli@st.com

Acknowledgements

We would like to take this opportunity to thank all contributors of this book, and all people working into the two involved research groups: Image Processing Lab (Catania University) and Advanced System Technology, Catania Lab (STMicroelectronics). A special thanks to Massimo Mancuso for having contributed with his extraordinary tenacity and competence to establish an imaging R&D group in Catania. We also thank prof. Giovanni Gallo for his invaluable help and support.

Fundamentals and HW/SW Partitioning

S. Battiato, G. Puglisi

Image Processing Lab, University of Catania, Italy.

A. Bruna, A. Capra, M. Guarnera

Advanced System Technology - Catania Lab, STMicroelectronics, Italy.

Abstract: The main goal of this Chapter is devoted to provide all the fundamental basis related to the involved technological issues relative to the single-sensor imaging devices. A rough understanding of the overall ingredients of a typical imaging pipeline is important also to consider the performance of any imaging devices, from low to high level, as the result of several components that run together to compose a complex system. The final image/video quality is the result of a certain number of design choices, that involve, in almost all cases, all aspects of the hardware and software technology. As briefly stated in the preface, the book aims to cover all aspects of algorithms and methods for the processing of digital images acquired by imaging consumer devices. More specifically, we will introduce the fundamental basis of specific processing into CFA (Color Filter Array) domain such as demosaicing, enhancement, denoising, compression together with ad-hoc matrixing, color balancing and exposure correction techniques devoted to preprocess input data coming from the sensor. We conclude the Chapter just including some related issues related to the intrinsic modularity of the pipeline together with a brief description of the hardware/software partitioning design phase.

1.1 The Simplest Imaging Pipeline

A typical imaging pipeline (see Fig.(1.1)) is composed by two functional modules (pre-acquisition and post-acquisition) where the data coming from the sensor in the CFA format are properly processed. The term pre-acquisition is referred to the stage in which the current input data coming from the sensor are analyzed just to collect statistics useful to set parameters for correct acquisition. In some cases several application can be present.

The initial data is composed by a matrix of data, coming from the sensor. For each pixel only a single chromatic value is acquired just using suitable CFA, typically arranged in the classic Bayer format. We omit all the details about optics and sensor capabilities that will be deeply treated in the next Chapter. Starting from the CFA data ad-hoc algorithms and methods can be used to obtain, at the end of the process, a compressed RGB version of the acquired scene. Some high-end devices allow the saving of the input data without applying any kind of processing, including compression, just providing as output an intermediate format, called "raw" format, where each pixel contains values very similar to those acquired by the sensor in the corresponding photosite. In the remaining cases, an imaging pipeline is needed to reconstruct (or recover) the missing data, maximizing whenever is possible, the related image quality. In the following Subsections we briefly summarize, with some examples, the typical (and mandatory) processing steps, just providing some initial overview of the relative algorithms that will be treated in more details in the rest of the book.

As depicted in Fig.(1.1) there could be a series of functional blocks devoted to im-

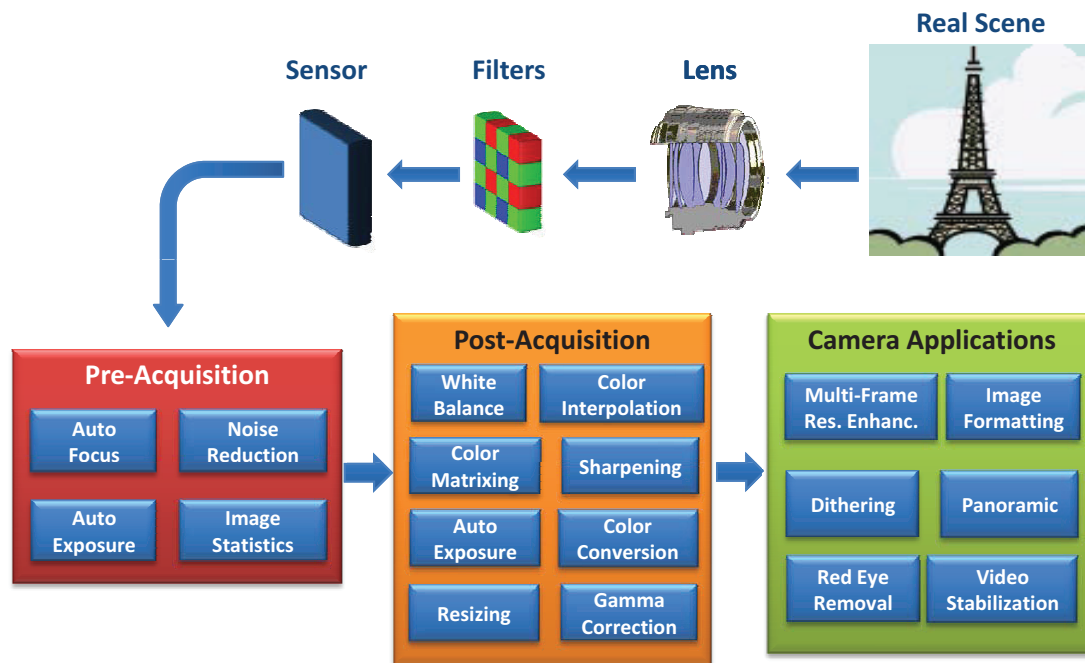


Figure 1.1 : Typical imaging pipeline. Data coming from the sensor (typically in Bayer format) are first analyzed to collect useful statistics for parameters setting (pre-acquisition) and then properly processed in order to obtain, at the end of the process, a compressed RGB image of the acquired scene (post-acquisition and camera applications).

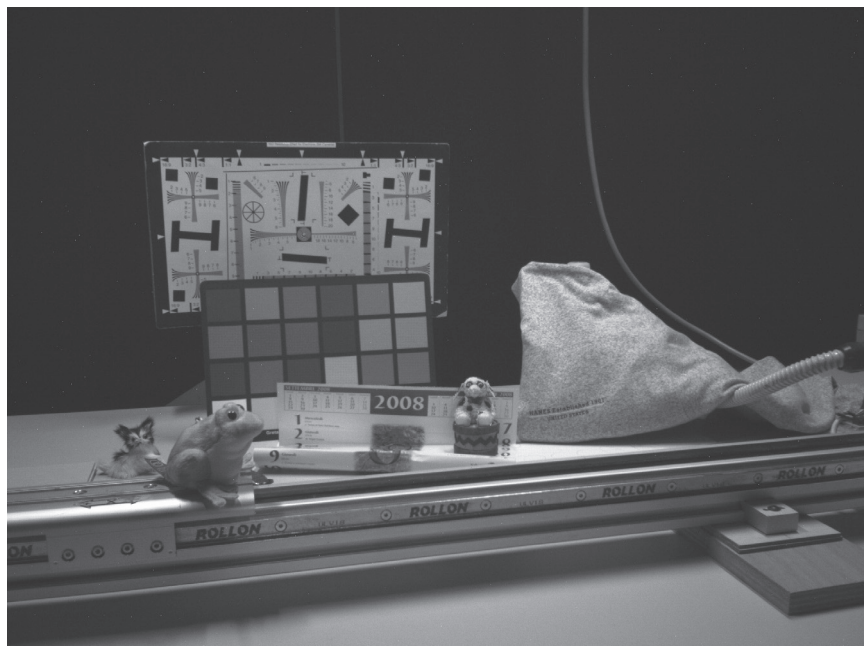
plements some specific camera applications: This functionalities are not mandatory and usually include solution for panoramic, resizing, red-eye removal, etc. Some of them could also require the multiple acquisition of the input scene at different exposure and/or focus settings (e.g., bracketing). An example of Bayer image, acquired by the monochromatic sensor, and the corresponding RGB image, obtained at the end of the pipeline, is shown in Fig.(1.2) and in Fig.(1.3).

Other related info can be found on [1], that is mainly devoted to cover aspects relative to optics and sensors, and [2] that addresses specific research challenges and recent trends.

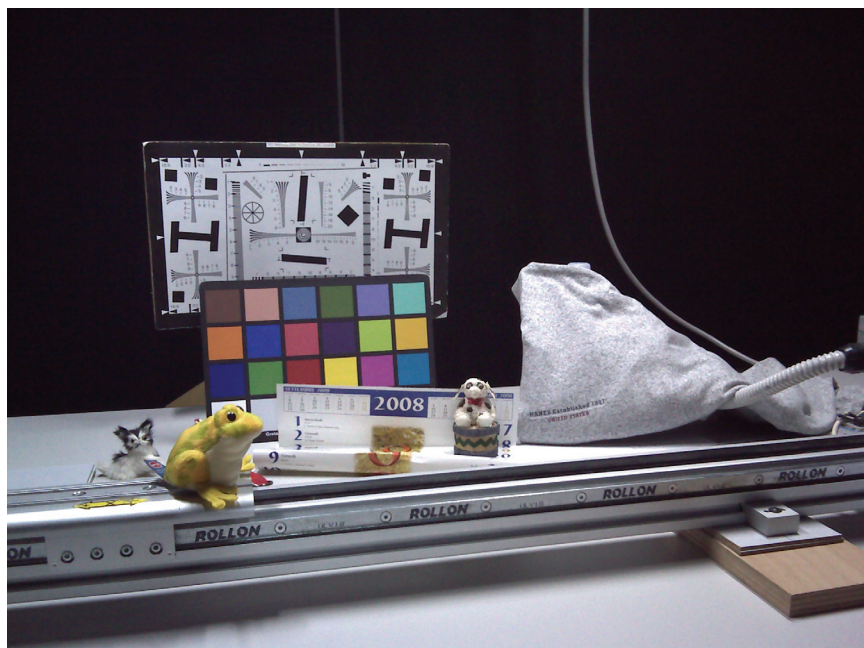
1.1.1 Exposure Setting

Like in old fashioned film cameras digital sensors need to be correctly exposed during acquisition. The pixel (picture element) is compound of an electronic device sensitive to the light (photo-diode or photo-transistor) which collects and translates incident photons (the electromagnetic element of the light) to electric signal. This signal is stored into an accumulation cell and, after an analog to digital conversion, represents the final pixel value (for detailed explanation see Chapter 2).

This basic light acquisition device has a few constraints: light sensitivity is fixed and



(a)



(b)

Figure 1.2 : An example of Bayer image (a) acquired by the monochromatic sensor and the corresponding RGB image (b) obtained at the end of the pipeline.

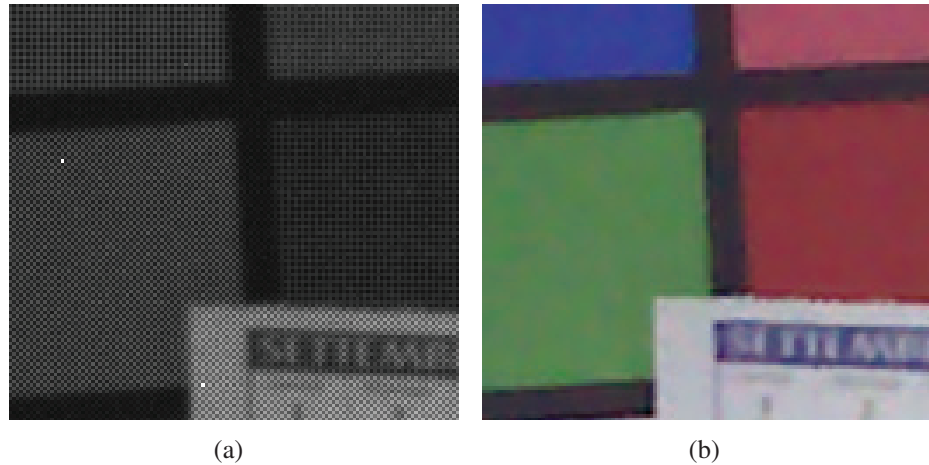


Figure 1.3 : An enlarged detail of Fig.(1.2) (a) acquired by the monochromatic sensor and the corresponding RGB image (b) obtained at the end of the pipeline.

it may be affected by noise (i.e., any kind of not actual information wrongly converted as useful information). Usually noise level is limited and not dangerous until the actual signal is adequate and significantly greater, i.e., high level of Signal to Noise Ratio (SNR).

To guarantee this fundamental principle each photosite (pixel) must be configured so that it acquires the correct level of light and thus the correct level of signal: varying the light intensity of the scene there must be a way to change the capability of the sensor to correctly and properly store in its cell the right level of light. This control is performed by the integration time. It represents the time during which the photo-element is acquiring and converting light into electrical charge. The lower the light intensity of the scene the higher the integration time. By changing this integration time a given scene digital acquisition can be under-exposed (too dark, too short integration time), over-exposed (too lit, too long integration time) or correctly exposed.

Two cases must be avoided or considered extreme cases: no accumulation, which corresponds to black, and over-accumulation (also known as saturation) which corresponds to extreme light or white. For actual black or white it is correct that the pixel assumes these values but they can also come out from a bad exposure (black from under-exposition and white from over-exposition). Also, there is no way to control integration time separately for each pixel of the sensor and this means that all the pixels of the sensor are exposed with the same integration time, although frame by frame it may change to adapt to variations of light which may occur in real scenes. Usually the integration time value is chosen so that the mean brightness of a picture is around the mid-range of the possible values (e.g., for a 8 bit per pixel image there are 256 different light levels and a correct exposed image has a mean brightness of 128).

Finally, it is not always possible to select an appropriate integration time for each scene. Too long or short integration time are not feasible because other problems may occurs and affect the SNR (for details see Chapters 2, 3 and 6). Also, integration time may be lower-limited by the framerate and/or by a safe value which aims to reduce motion

blur effects. Motion blur is caused by long integration time and moving objects in the acquisition scene or hand-shaking. It is typical in low-light acquisitions and for this reason often a flash is used in such a situation.

Each time a lower threshold limits the integration time the only way to properly read the minimum accumulated information of the cell is to use a multiplicative gain to amplify the information to a usable value.

In summary, a good exposure control module is compound of:

- an appropriate module which estimates the light intensity of the scene and properly settles the correspondent integration time, avoiding under-exposition or over-exposition;
- an appropriate gain control which furnishes support and compensates the limits of the integration time; when selecting a proper balance between integration time and gains priority goes to the former;
- a method to identify actual black and white regions, assumed that the rest of the identification and proper compensation is demanded to following modules in the image generation pipeline (see Chapter 4 for additional details);
- a loop with other modules which apply additional gains to the signals (like AWB, see Chapter 5);
- and additional optional module to control and avoid motion blur; usually in literature this methodology goes with the name of AutoISO.

1.1.2 White Balance

One of the most challenging processes that affects perceived image quality in a digital camera is the correct color reproduction. Human visual system is able to remove color casts: an object appears to our eyes with the same color under different illuminant conditions. On the contrary the sensor simply acquires raw data and is not able to cope with real scene illumination variability. For instance a white paper in outdoor or indoor environment can be recorded by the sensor with bluish or reddish colors.

In order to cope with these problems a lot of techniques have been developed. High end cameras typically provide a variety of presets related to the most common light sources (tungsten, fluorescent, daylight, flash, etc.). Moreover white balance parameters can be set, for future photos, taking a picture of a known gray reference under the same illumination source (custom white balance).

All the techniques above described need a close interaction with the user in order to properly work. On the contrary auto white balance techniques try to guess the correct illumination properties and remove color casts without user interaction. These techniques, based on strong assumptions on scene reflectance distribution, have been also implemented in low cost devices (e.g., smart phone) and will be in depth described later in Chapter 5.

1.1.3 Noise Reduction

The perceived image quality is deeply influenced by image noise (named by analogy with unwanted sound). These unwanted fluctuations, if not properly managed, heavily degrade image quality. Different noise sources, with different characteristics, are superimposed to the image signal: photon shot noise, dark current noise, readout noise, reset noise, quantization noise, etc.

Although many efforts have been done by manufacturers to reduce the presence of noise in imaging devices it is still present and can be considered unavoidable in critical situations. For instance, low light conditions together with low integration time, produce very low SNR (signal to noise ratio), very few photon were captured, making really difficult obtaining pleasing photos. This physical limit does not depend only on the sensor characteristics but it is strictly related to the nature of light. Moreover the increasing of the number of pixels and the limited size of the embedded devices, implying the decreasing of the pixel size, produces further problems. Small pixels, acquires less photons with respect to larger pixels. Less useful signal implies then noisier picture.

In order to cope with these problems smart filters must be designed. These filters must be able to estimate image noise characteristics (e.g., mean and standard deviation if a Gaussian model is used), and then remove unwanted noise without affecting image details.

Finally, it should be noted that noise reduction can be performed during the various stages of the pipeline. Some approaches works on RGB images, others directly on Bayer data. The latter typically provides some advantages (demosaicing step typically introduces nonlinearities that make difficult noise reduction). Further details about noise reduction algorithms will be provided in Chapter 6.

1.1.4 Demosaicing

Digital cameras, in order to reduce costs and complexity, acquire images by means of a monochromatic sensor covered by a CFA (color filter array). A lot of CFA have been developed but the most common is the Bayer pattern. This simple CFA, taking into account human visual system characteristics (human eyes are more sensitive to green with respect to the other primary colors), contains twice as many green as red or blue sensors. Some spatially undersampled color channels (three in the Bayer pattern) are then provided by the sensor and the full color information is reconstructed by color interpolation algorithms (demosaicing). Demosaicing is a very critical task. A lot of annoying artifacts that heavily degrade picture quality can be generated in this step: zipper effect, false color, moiré effect, etc. Simple intra-channel interpolation algorithms (e.g., bilinear, bicubic) cannot be then applied and more advanced solutions (inter-channel), both spatial and frequency domain based, have been developed. In embedded devices the complexity of these algorithms must be pretty low. Demosaicing approaches are not always able to completely eliminate false colors and zipper effects, thus imaging pipelines often include a post-processing module, with the aim of removing residual artifacts. Further details about demosaicing algorithms will be provided in Chapter 7.

1.1.5 Color Matrixing

The Color Matrix sub-system, also known as Color Calibration, aims to convert the color response of the acquisition device to a standard color space. Usually the standard RGB (sRGB) color space is used, according to the ITU-R BT.709 directive. This transformation is needed since the spectral sensitivity function of the sensors does not match with the desired color space. The correction is performed usually according to the formula:

$$\mathbf{RGB}_{out} = \mathbf{A} \cdot \mathbf{RGB}_{in} \quad (1.1)$$

where $\mathbf{A}$ is a 3-by-3 matrix, $\mathbf{RGB}_{in}$ and $\mathbf{RGB}_{out}$ the image before and after color matrixing. The matrix coefficients are not obtained using the effective response. Usually they are retrieved using optimization methods with real acquisitions. Moreover the constraint of the white point preservation is usually used. It corresponds to the following constraint (as better detailed in Chapter 5):

$$\sum_{j=1}^3 A_{(i,j)} = 1, \forall i \in \{1, 2, 3\} \quad (1.2)$$

1.1.6 Image Formatting

The data acquired by the sensor have to be processed by the coprocessor or the host microprocessor, so both the systems must share the same communication protocol and data format. Moreover, at the end of the image generation pipeline the image must be coded in a standard format in order to be read by any external device. Usually the sensor provides the acquired image in the Bayer format. In the past the Bayer data were stored and transmitted using proprietary format and protocol. Such solution has the drawback that every customer had to design the same proprietary interface to manage the sensor data. In the latest years the main companies making, buying or specifying camera modules proposed a new standard called Standard Mobile Imaging Architecture (SMIA). It allows interconnecting sensors and hosts of different vendors.

Concerning the output of the coprocessor, several standard formats are available. For the still images the most frequently used are the Joint Picture Expert Group (JPEG) with a lossy compression, the Targa Interchange Format (TIF) with a lossless compression. In the top level cameras the output of the sensors can also be stored directly. In this case usually a proprietary file format is used (e.g., the Nikon Electronic Image Format (NEF), the Canon RAW File Format (CRW), etc.). For videos the most used are Motion JPEG, MPEG-4, H.263 and H.264 standards.

In Chapter 11 the main data formats will be presented. Moreover some techniques concerning the compression factor control and the error concealment will be introduced. The compression factor control aims to obtain the file size as close as possible to a target value whereas the error concealment aims to handle errors in the bit-stream trying to recover the missing information.

1.2 HW/SW Partitioning

Cameras embedded in mobile phones are now becoming a commodity supporting applications like capturing and transmission of still images as well as video clips (Multimedia Messaging Services). With the increase of network bandwidth (e.g., 3G UMTS) real time mobile video links will become feasible, enabling new applications like mobile video telephony and video chat. It has to be noted, that the ease of use of these applications is of high importance as this is expected to be a crucial requirement for market acceptance of such new services. Thereby not only quality issues like frame and image stabilization are to be focused but also the user comfort. The automatic detection and tracking of the user's head is such an example, which helps to keep one's face in view of the camera during a mobile video telephone conference. But the processing units in imaging devices should be low-cost, low-power and, at the same time, suitable of supporting the above mentioned mobile communication applications. In order to satisfy cost and performance requirements, imaging device systems are generally implemented with a combination of different components, from custom designed accelerators to standard processors. These components can vary in their area, speed, methodology to program, and the system functionality is partitioned amongst the components to best utilize this tradeoff. However, for performance critical designs, it is not sufficient to only implement the critical sections as custom-designed high-performance hardware, but it is also necessary to pipeline the system at several levels of granularity. The custom designed accelerators can be implemented by using Reconfigurable hardware devices, such as Field Programmable Gate Arrays (FPGAs). The HW/SW partitioning (i.e., the definition of an architecture where the algorithms are smartly split as hardware accelerators and software modules) is not as straightforward as designing either software or hardware, since the application is intrinsically a hardware/software co-design. For instance, while an application implemented on an FPGA can be one to two orders of magnitude faster than the application implemented in software, processing in hardware incurs additional costs that are not required for software. Some of these costs are hardware initialization costs, extra processing steps for easy processing of the border cases, and communication of the image to and from the reconfigurable device. The runtime of image processing applications varies with image size, so processing small images on an FPGA might not be efficient due to the additional overhead. The imaging accelerators are often designed to create data-paths that are capable to process several image pixels concurrently. For the definition of these data path can be used some well-known design approaches like:

- SIMD parallelism. Typically the data-path processes N pixels in parallel or, for some binary operation, $8 \times N$ pixels. This type of processing is well known from multi-media extensions used in general-purpose CPUs.
- Deeper arithmetic pipelines. These enable the encoding and execution of complex arithmetic operations with a single microinstruction.

Bibliography

- [1] J. Nakamura, *Image Sensors and Signal Processing for Digital Still Cameras*. CRC Tailor & Francis, 2006.
- [2] R. Lukac, *Single-Sensor Imaging: Methods and Applications for Digital Cameras*. CRC Press, 2008.

Notions about Optics and Sensors

A. Bruna, A. Capra, M. Guarnera, G. Messina

Advanced System Technology - Catania Lab, STMicroelectronics, Italy.

Abstract: This Chapter gives information about the optics and the sensor in an image acquisition system. Optic is the first stage of the image acquisition system and is composed by one or more lenses aiming to concentrate the light in the physical sensor obtaining an in focus image. It is one of the most expensive parts of an imaging system. In mobile cameras usually a lens is compound of a system of plastic and glass lenses stacked together, while in single lens reflex cameras (SLR) a group of several glass lenses system is employed to reduce image artifacts. In this Chapter an overview of the lenses will be discussed. Moreover, some typical artifacts will be introduced (e.g., cross talk and chromatic aberration). The sensor is the part that converts the optical image (light) to an electric signal. There are several kinds of sensors depending on the technology (CCD - Charge Coupled Device and CMOS - Complementary MetalOxideSemiconductor), on the color filter array (Bayer , Foveon, 3CCD, Panchromatic), on the transducer function (LDR - low dynamic range, WDR - wide dynamic range and HDR - high dynamic range sensors).

2.1 Introduction

In Fig.(2.1) is reported the classical "yellow duck" system that describes the entire process from the real world, the system acquisition, the digital processing up to the final digitally compressed image. The first element, the lens and its working within the complete system, will be deeply studied in Chapter 4. This Chapter is mainly devoted to the explanation of the parts inside the box *Image Sensor*, the silicon sensor, its surface, its technology and its basic element: the pixel (picture element).

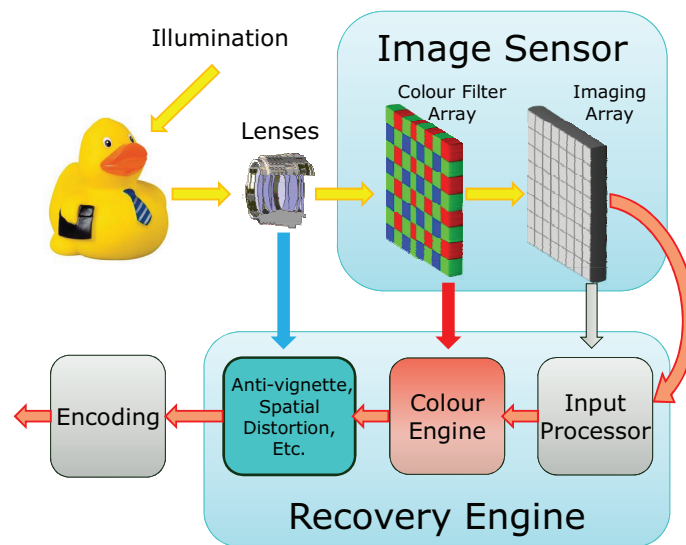


Figure 2.1 : The yellow duck system: from the real world to the final digital picture.

Anatomy of the Active Pixel Sensor Photodiode

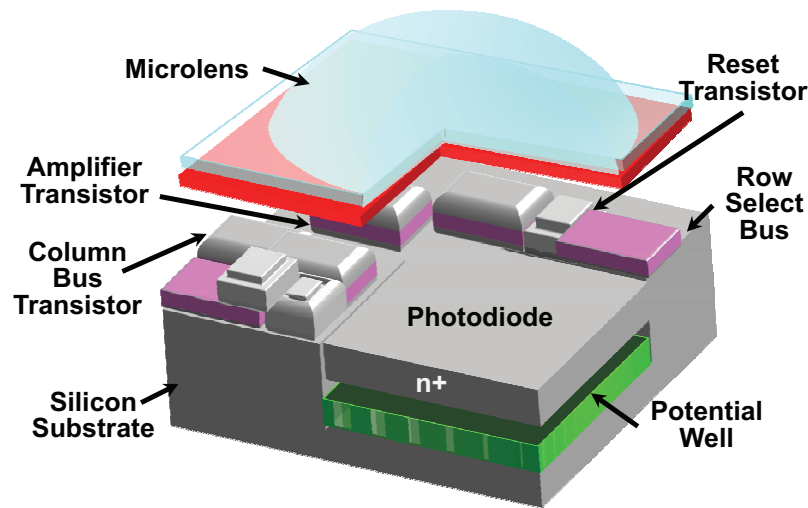


Figure 2.2 : Example of a pixel structure.

In the following Subsections will be explained the overall description of the sensor system, the different sensor technologies and the color patterns used in the real devices.

2.2 The Sensor System

In Fig.(2.2)¹ is reported a classical structure of a pixel inside a sensor. All the related compounding elements are visible: micro-lens, electronic part and photo-sensible area. In the following Subsections all these elements are analyzed, describing their strongness and weakness and their contribute in the resulting final image.

2.2.1 Microlenses

One of the most important aspect of a imaging sensor is the sensitivity to capture the light (light sensitivity). As we will see in detail in Sec.2.3 it mainly depends by the design of the photo-sensible area, Fig.(2.2), and, for equal design and technology, by the size of this area. The ratio between the photosensitive area and the pixel size is indicated as fill factor (it may vary from 30% to 100%). When this fill factor is too low and needs to be improved the microlenses are used. They are individual lenses deposited on the surface of each pixel to focus light on the photosensitive area. Microlenses can boost effective fill factor (i.e., the ratio of the active refracting area to the total area occupied by the lens array) to approximately 70%, improving sensitivity (but not charge capacity)

¹Image source: http://digitalcontentproducer.com/hdhdv/depth/cmos_tech_hdv_10092006

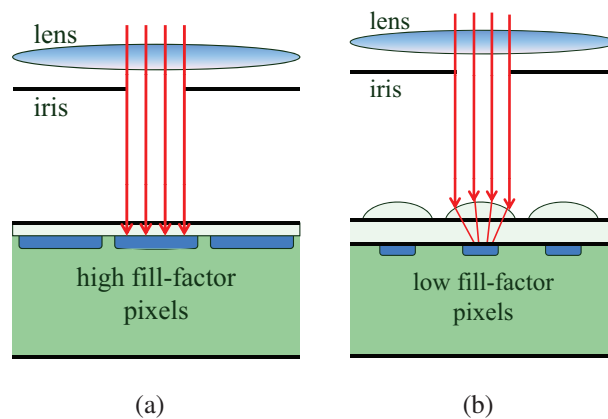


Figure 2.3 : Strategies to compensate different pixels fill factor: **(a)** Pixel with good fill factor, all the entering light is captured; **(b)** Pixel with poor fill factor, a micro-lens is used to converge the entering light to the useful region.

considerably. In Fig.(2.3)² is illustrated the difference of a pixel with and without micro-lens.

2.2.2 Lenses and Microlenses Aberrations

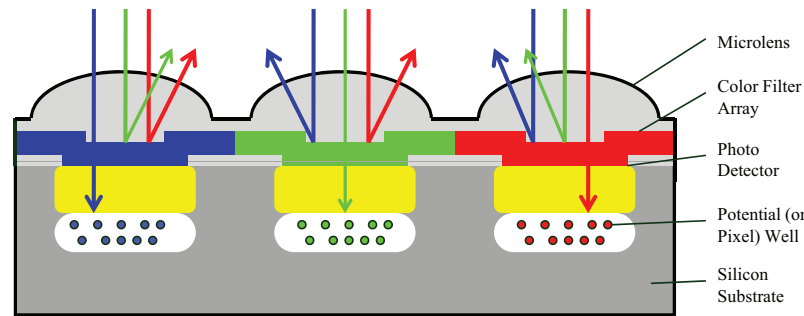
Crosstalk

In an ideal pixel charging [2], all the photons intersecting at any angle a filter element in the CFA are color filtered and accumulated in the photodetector under the filter element as shown in Fig.(2.4(a)). Optical crosstalk results when a photon, intersecting at an angle with a filter element in the CFA, enters the adjacent pixel's photodetector (photodiode) rather than the photodetector under the filter element. This can contaminate the adjacent pixel's charge packet, Fig.(2.4(b)). Moreover there may be also electrical crosstalk, which happens when photons passing through the red filter travel further into the silicon before generating electrons. This leads to a non-uniform response to the different colors, a loss of charge into the substrate and electrons wandering into the wrong pixel well as shown in Fig.(2.4(c)).

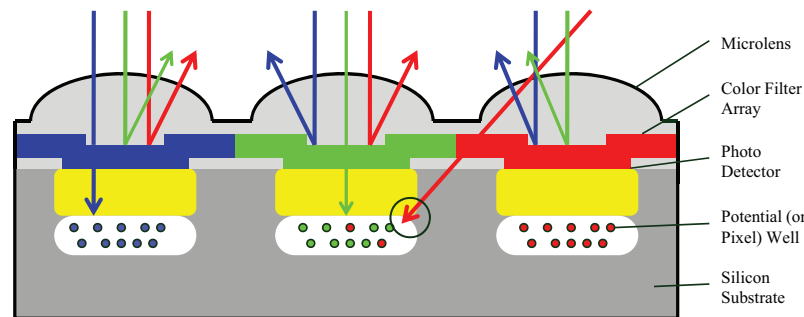
Chromatic Aberration

Chromatic Aberration (CA) is the term for an imaging system placing incorrect colors in at least some locations within the image. CA is caused by a lens having different refractive indices to different wavelengths of light. This implies the color channels to focus differently, and hence CA creates visible color fringes or colored blurs. An example of CA is shown in Fig.(2.5). Such aberration is typically diffused in the entire image, but

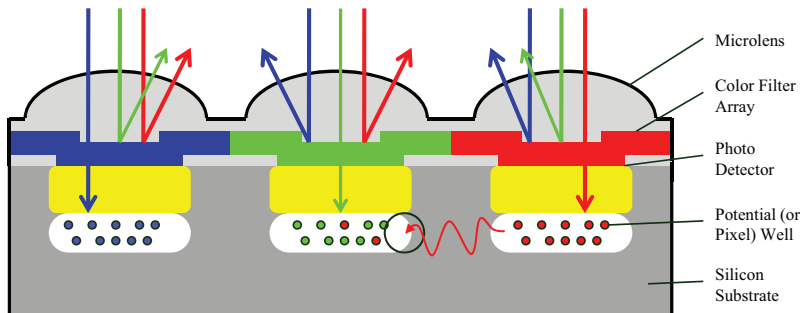
²Image source: [1]



(a)



(b)



(c)

Figure 2.4 : Different types of crosstalk: **(a)** In this case all the entering light is correctly captured by the detector, there is not crosstalk effect; **(b)** Optical crosstalk: some filtered light is not captured by the respective detector but goes to an adjacent detector, in this case red light is collected in the green collector; **(c)** Electrical crosstalk: some photons are dispersed in the silicon and collected by a different detector, in this case a red photon travelling in the silicon is collected by the adjacent green detector.



Figure 2.5 : Particular of an image affected by CA.

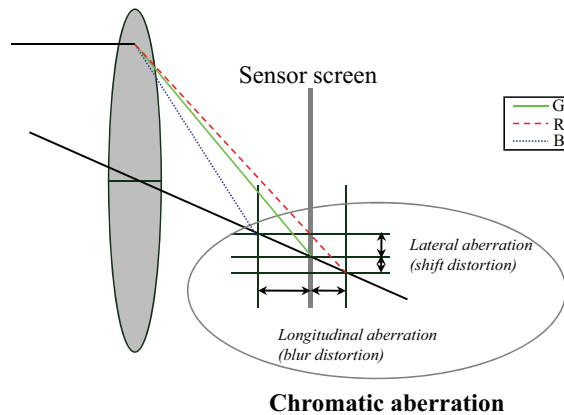


Figure 2.6 : Longitudinal and lateral aberrations.

it is mostly observed near high contrast edges. CA is usually classified in two types (see Fig.(2.6)):

- Longitudinal aberration;
- Lateral aberration (also known as aberration of magnification).

Longitudinal aberration causes loss of focus in the image plane according to the wavelength, and lateral aberration occurs due to the shift in the focus. The former causes blur distortion and the latter causes shift distortion of the color channels. Imaging devices usually focus the image plane based on the Green channel, as shown in Fig.(2.7), because it is a dominant factor for brightness. Therefore the Green channel exhibits little CA as compared to the other two color channels.

The two main methods for correcting chromatic aberration are:

- Lens design for improved optics;

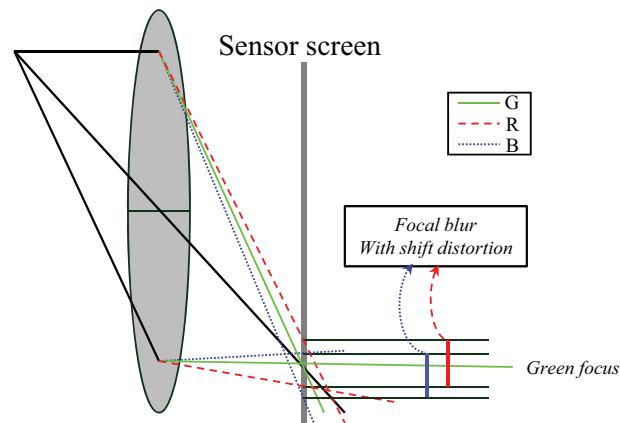


Figure 2.7 : Distortions of the R and B channels.

- Application of a signal processing method to the acquired image [3,4].

The first approach is generally expensive and difficult to apply. On the other hand, signal processing algorithms can correct this defect in a cheap and effective manner.

Simplest techniques measure the chromatic aberration of magnification on standard patterns, such as chessboard patterns. An example of pattern used for calibration is shown in Fig.(2.8) (left side). A second lower density chessboard pattern (test image), shown in Fig.(2.8) (right side), is used for independent validation of the results. In particular as discussed in [3], the reference information on the standard pattern without chromatic aberration is prepared in advance, then the standard pattern is actually photographed and the color shift is detected by comparing this data with the reference information .

Conventional image processing algorithms for correcting chromatic aberration usually make use of supplementary information about the optical system to calculate the color shift due to the chromatic aberration of magnification. The technique proposed in [4], tries to perform the correction of the lateral chromatic aberration by introducing differential geometric distortion into the color components of the image object. This distortion is used to substantially cancel the differences in image heights of the red, green and blue color components of the virtual image formed by the lens element. This is accomplished by determining one or more color-specific distortion functions (taking into account the different lateral magnification parameters of lens element for red, green and blue image planes) which are applied to the geometry of one or more color components. Differential distortion functions may be applied to the red and blue color components to align them with the green color component.

Purple Fringing

Although chromatic aberrations can be purple in color under certain circumstances, "Purple Fringing" usually refers to a typical imaging device phenomenon that is caused by the

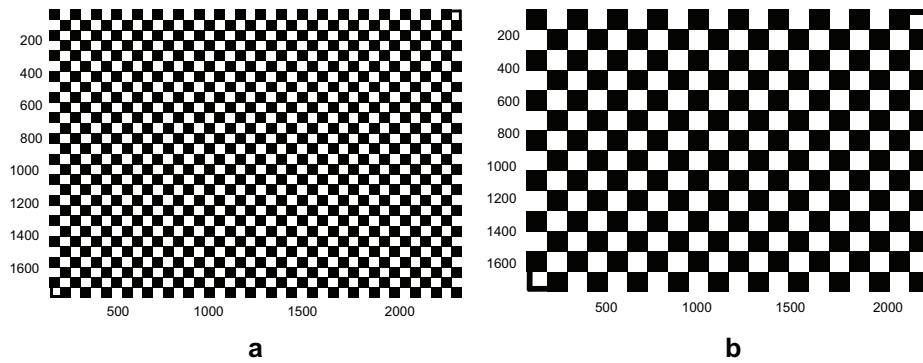


Figure 2.8 : Chessboard patterns used for calibration (left) and testing (right).



Figure 2.9 : Particular of an image affected by PF.

microlenses. It can be considered as "chromatic aberration at microlens level". As a consequence, purple fringing is visible throughout the image frame. An example of image affected by purple fringing is shown in Fig.(2.9). Edges of contrasty subjects suffer most, especially if the light comes from behind them. Blooming tends to increase the visibility of purple fringing. In fact imaging devices usually meter the exposure of the scene so that in the brightest regions as much charge is accumulated as possible without overflowing the potential well, so giving the best dynamic range and a response that depends directly on the amount of light striking the photodiode. The exposure control normally works well but the metering of the scene is done in some average sense. If the scene contains small regions where the light is extremely bright compared to the scene average, these bright regions lead to a very large amount of light being incident on the sensor. This light causes charge to build up in the sensor until a sufficiently large amount accumulates that it begins to leak out of the sensor well into the surrounding sensors (blooming effect).

When charge leaks from one photodiode well into surrounding photodiode wells, the result is a spuriously higher signal in the surroundings. That spurious signal will be particularly noticeable if the surrounding sensors should be producing no signal because

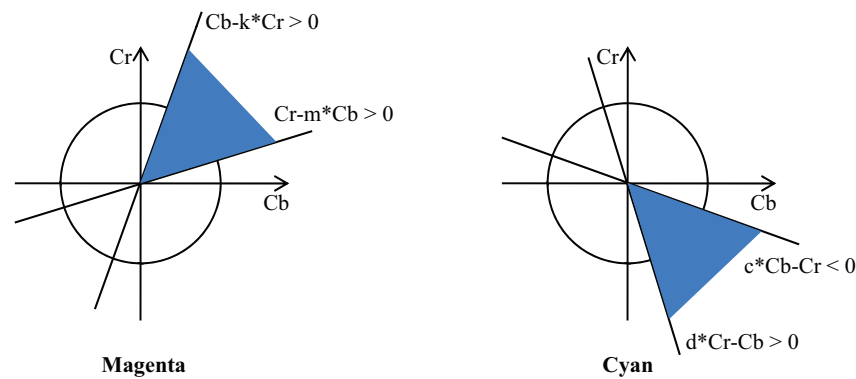


Figure 2.10 : Locus of purple in CMY color space.

the scene is dark in these regions. Is thus easier to see the effect of blooming most strongly at sharp transitions from light to dark. We have already seen that lens aberration will cause the blue and red components of bright white light to appear at incorrect sensor positions. Charge leakage magnifies this effect by spreading the sensor response further from the true position, and the positional error is additionally potentiated by demosaicing

Most of the state of the art algorithms devoted to correct the purple fringe effect construct a binary map, discriminating overexposed regions from non-overexposed ones. They detect purple regions close to overexposed areas and desaturate them. Some of these techniques try to prevent an excessive desaturation, according to the saturation of the surrounding pixels, to provide high quality images. Some other techniques try to identify the occurrence of purple fringing by detecting magenta and/or cyan color ranges in the $Cr Cb$ domain, as shown in Fig.(2.10).

An interesting technique [5] for removing purple fringing assumes purple fringe appearing principally around an overexposed highlight, then the false color degree of each pixel is calculated, and appropriate correction is performed corresponding to these false color degrees. A pixel which satisfies the following conditions:

- an overexposed pixel exists around the noticed pixel;
- the noticed pixel exhibits a purple color;
- the noticed pixel has a high saturation.

is determined as a purple fringe pixel. The correction process can be executed by averaging neighboring pixels which do not exhibit a false color.

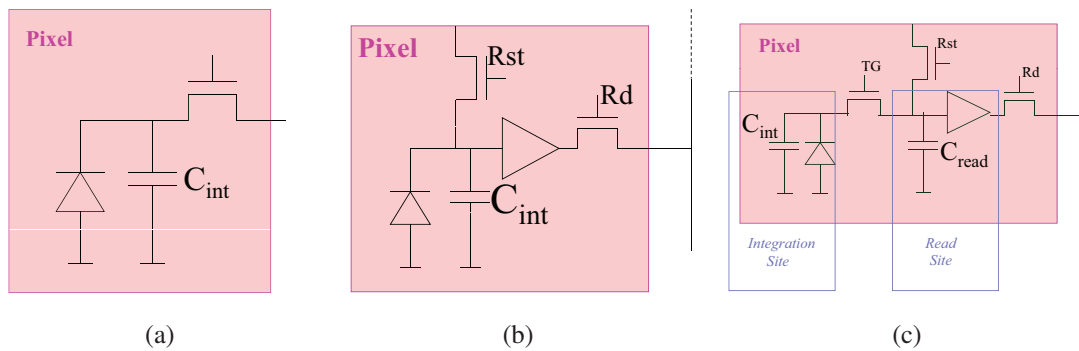


Figure 2.11 : Structure of a pixel and its evolution; **(a)** pixel in its basic structure: photodiode and accumulation cell; **(b)** active pixel, an on site amplification has been added; **(c)** active pixel in 4-T configuration: accumulation and reading phases are separated.

2.3 The Sensor Technology

Imaging sensor functionality for digital camera is the same of a film for old camera. It is used to convert light entering from the lens system to electrical signal which is properly processed and stored at the end of the pipeline in a standard image format file. The basic structure of a pixel is reported in Fig.(2.11(a)): it is compound of a photodiode to convert light to an electrical signal and an accumulation cell (capacitor) and a switch. This configuration is known as passive pixel. The advantage of this configuration is the maximization of the fill factor.

Adding a signal amplification to the pixel, it is possible to have an active pixel, as shown in Fig.(2.11(b)). The amplification system allows to improve the signal level during the reading process, thus improving the SNR. It is also known as 3-T (3 transistors) configuration. The charge to voltage conversion within the pixel improves output impedance but it causes a lower fill factor, because now part of the pixel surface is dedicated to two switches and an amplifier. A sensor compound of such a kind of pixels goes by the name of Active Pixel Sensor (APS) [6].

Nevertheless this configuration has the accumulation and the reading stages not separated and thus, during the reading process, noise coming from the accumulation part is amplified and transmitted. For this reason a widely used configuration is the 4-T configuration, as shown in Fig.(2.11(c)). In this case accumulation and reading are separated, improving the final SNR.

2.3.1 CCD and CMOS Sensors

First imaging sensors were introduced around 1970 and were based on CCD (Charge Coupled Device) technology. As for CMOS sensor they are silicon based and share similar materials. However CCD sensors have been specifically designed to capture images and have 40 years of development. Until a few years ago they were the state of the art as