A Unified Framework for the Consumer-Grade Image Pipeline

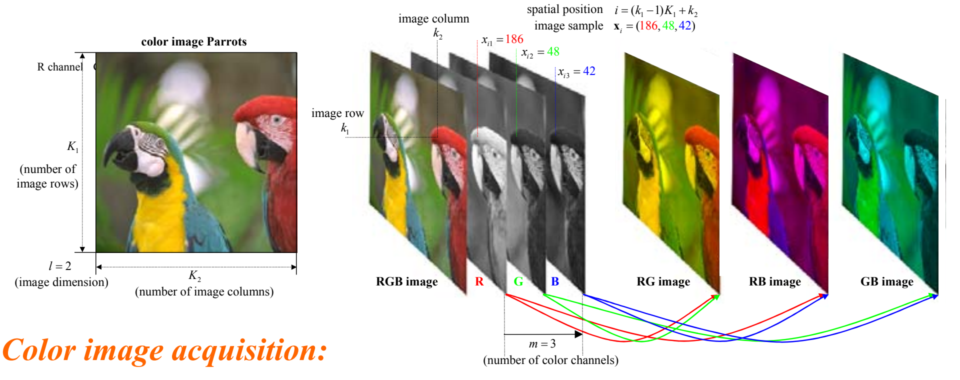
Konstantinos N. Plataniotis
University of Toronto
kostas@dsp.utoronto.ca
www.dsp.utoronto.ca

Common work with **Rastislav Lukac**

Outline

- *The problem*
- *Background*
- *Single Sensor Imaging: Challenges & Opportunities*
- *Performance issues*
- *Conclusions*

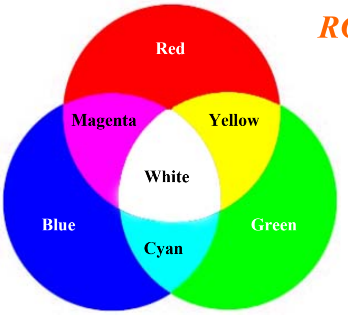
Digital color imaging



Color image acquisition:

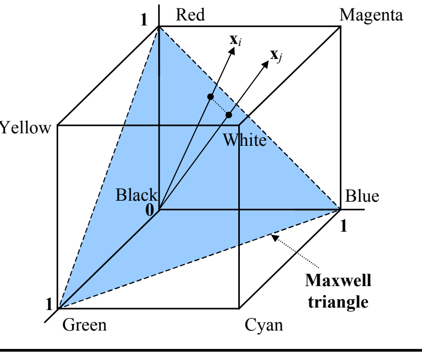
- digital cameras - most popular and widely used
- scanners
- synthetic (e.g. gray-scale image coloration)

Focusing on the color pixel level



RGB (sRGB) color space:

- commonly used for acquisition, storage, and displaying purposes
- additive concept of color composition

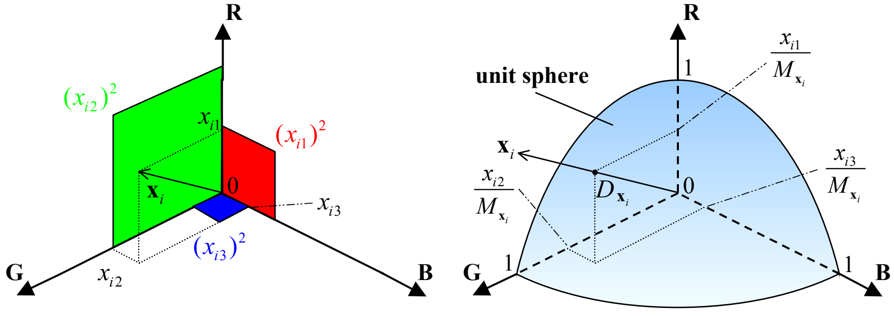


- RGB color pixel is the vector in a three-dimensional (RGB) color space
- vector components are the intensities measured in RGB color channels

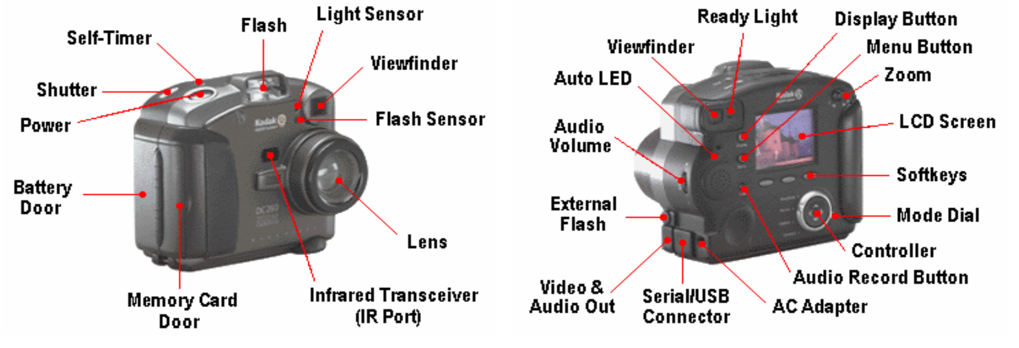
Color imaging basics

Color vector: uniquely characterized by its

- magnitude (length) $$M_{\mathbf{x}_i} = \|\mathbf{x}_i\| = \sqrt{(x_{i1})^2 + (x_{i2})^2 + (x_{i3})^2}$$
- direction (orientation) $$D_{\mathbf{x}_i} = \left(\frac{x_{i1}}{M_{\mathbf{x}_i}}, \frac{x_{i2}}{M_{\mathbf{x}_i}}, \frac{x_{i3}}{M_{\mathbf{x}_i}}\right); \quad \|D_{\mathbf{x}_i}\| = 1$$



Camera: End-user's point of view

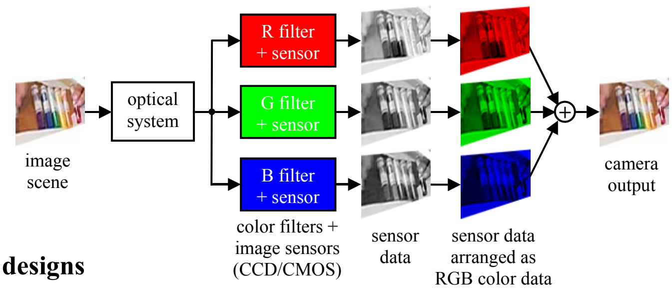


Focus on effectiveness: functionality vs. cost

- **optics (optical zoom), digital zoom, memory, battery, etc.**
- **multimedia acquisition, processing & transmission (image, audio and video)**

Three-sensor imaging

- "Sensor" : a monochromatic device; most expensive component of the digital camera (10% to 25% of the total cost)
  - charge-coupled device (CCD)
  - complementary metal oxide semiconductor (CMOS) sensor

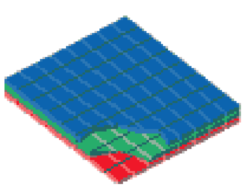
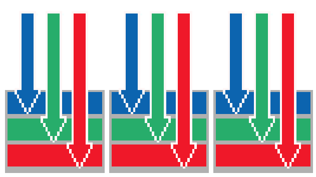
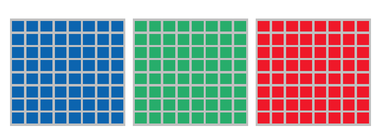


- professional designs
- each sensor corresponds to a particular color channel
- spectrally selective filters
- expensive solution

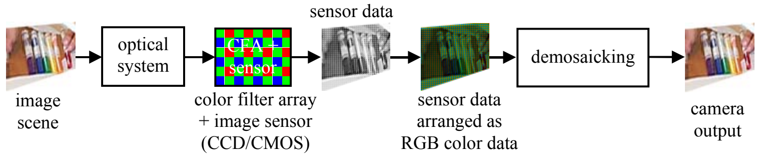
X3 technology-based single-sensor imaging

Layered (three-layer) silicon sensor

- new technology - expensive solution for professional devices (medical & science applications)
- directly captures RGB light at each spatial location in an image during a single exposure
- takes advantage of the natural light absorbing characteristics of silicon
- color filters are stacked vertically and ordered according to the energy of the photons absorbed by silicon

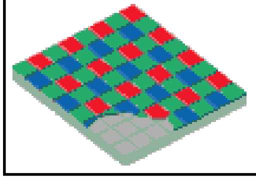




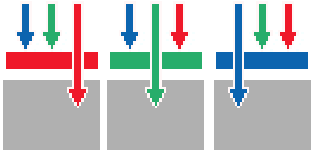
Single-sensor imaging

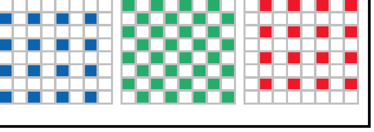


Color filter array (CFA)

- generates a 2-D array or mosaic of color components
- produced CFA (sensor) image is a gray-scale image
- full-color image is obtained through digital processing







Color filter array (CFA) design

Key factors in CFA design

- immunity to color artifacts and color moiré
- cost-effective image reconstruction
- reaction of the pattern to image sensor imperfections
- immunity to optical/electrical cross talk between neighboring pixels

Color systems used in CFA design

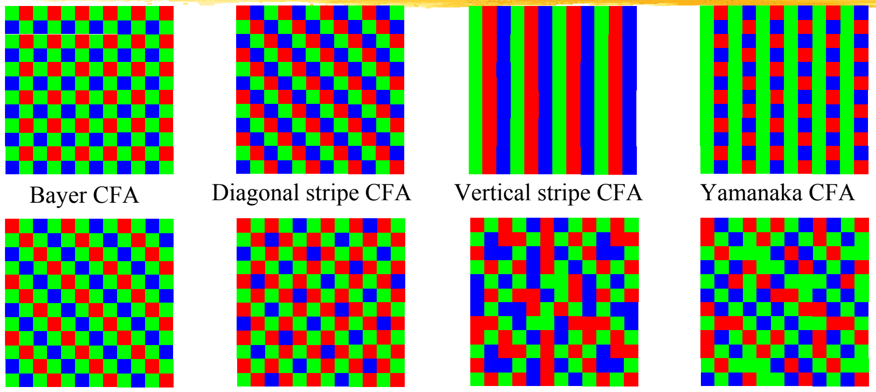
i) tri-stimulus (RGB, YMC) systems - RGB is most widely used

ii) mixed primary/complementary colors (e.g. MGCY pattern)

iii) four and more color systems (white and/or colors with shifted spectral sensitivity) -

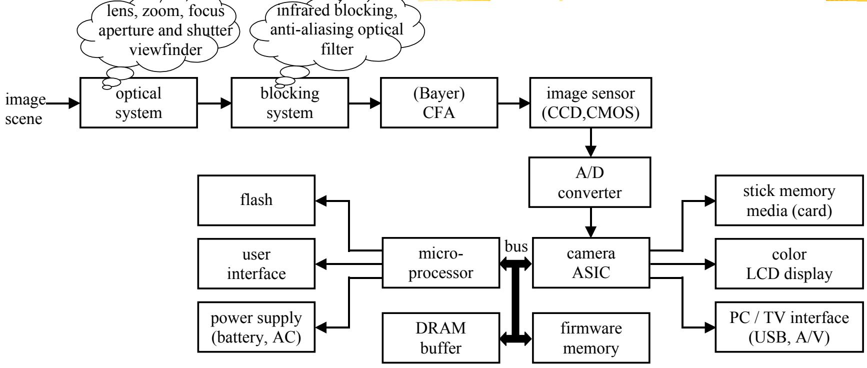
• CFAs in ii) and iii) may produce more accurate hue gamut, but they limit the useful range of the darker colors

Common RGB-based CFAs



- **Bayer CFA is widely used (good performance, cost-effective color reconstruction)**

Single-sensor camera architecture



- **DRAM buffer temporally stores the digital data from the A/D converter**
- **DRAM then passes data to the application-specific integrated circuit (ASIC)**
- **digital data processing, such as demosaicking and image resizing, is realized in both ASIC and microprocessor**

Camera image processing

✓*Processing*

- demosaicking (spectral interpolation)
- demosaicked image postprocessing (color image enhancement)
- camera image zooming (spatial interpolation in CFA or full-color domain)

Compression

- lossy (or near lossless) vs. lossless compression
- CFA image compression vs. demosaicked image compression

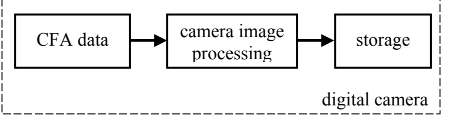
Data management

- camera (CFA) image indexing → connection to image retrieval

Implementation

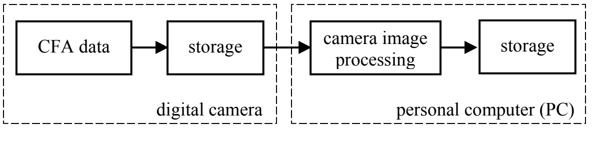
Conventional digital camera

- real-time constraints (computational simplicity requirements)



Using a companion personal computer (PC)

- PC interfaces with the digital camera which stores the images in the raw CFA format
- allows for the utilization of sophisticated solutions



Camera processing operations

Considering the spectral image characteristics

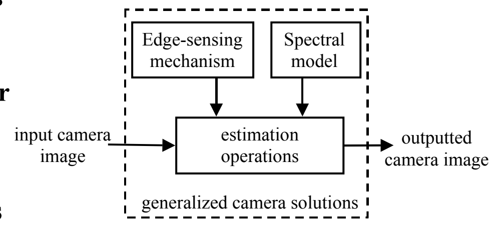
- component-wise (marginal) processing (component → component)
- spectral model-based processing (vector → component)
- vector processing (vector → vector)

Considering the image content (structure)

- non-adaptive processing
- (data) adaptive processing

Practical solutions

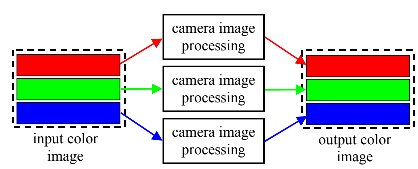
- spectral model used to eliminate color shifts and artifacts
- edge-sensing mechanism used to eliminate edge-blurring and to produce sharply-looking fine details



Considering the spectral characteristics

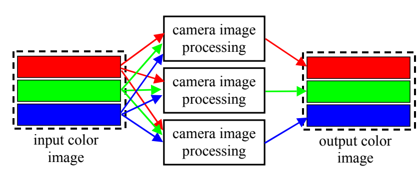
Component-wise processing

- each color plane processed separately
- omission of the spectral information results in color shifts and artifacts



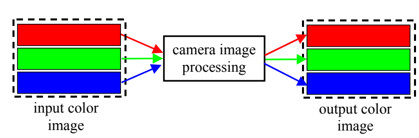
Spectral model based processing

- essential spectral information utilized during processing
- computationally very efficient - most widely used in camera image processing



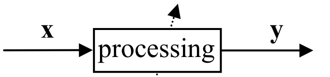
Vector processing

- image pixels are processed as vectors
- computationally expensive



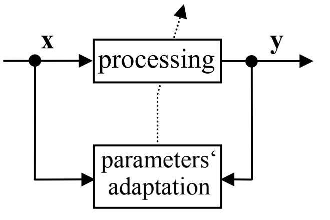
Considering the image content

Non-adaptive processing



- no data-adaptive control
- often reduces to linear processing - easy to implement
- performance limitations (image blurring)

Adaptive processing



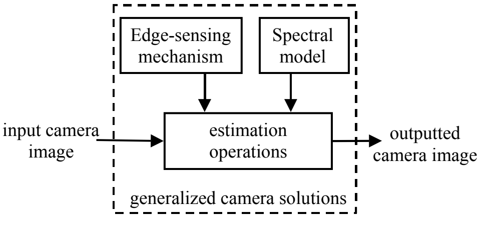
- edge-sensing weights used to follow structural content
- nonlinear processing
- enhanced performance, sharply looking images

Data-adaptive processing

Construction • using spatial, structural, and spectral characteristics

$$\mathbf{x}_{(p,q)} = \sum_{(i,j)\in\zeta} \{ w'_{(i,j)} \Psi(\mathbf{x}_{(i,j)}, \mathbf{x}_{(p,q)}) \}$$

$$w'_{(i,j)} = w_{(i,j)} / \sum_{(i,j)\in\zeta} w_{(i,j)}$$



Spatial characteristics

- local neighborhood area ζ

Structural characteristics

- edge-sensing mechanism λ

$$\lambda(z) \to \{ w_{(i,j)}, (i,j) \in \zeta \}$$

z denotes the CFA image

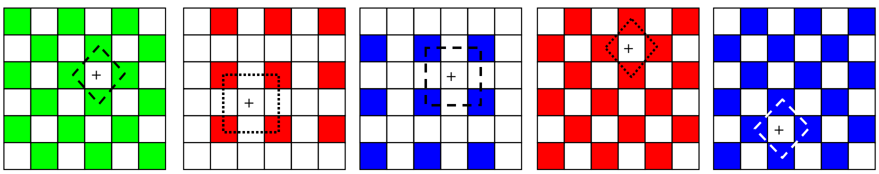
Spectral characteristics

- spectral model Ψ

Local neighborhood area

Features

- approximation using a shape mask ζ
- shape and size of ζ vary depending on the CFA used and processing task (demosaicking, resizing, etc.)
- shape masks widely used in the demosaicking process:



(a) (b) (c) (d) (e)

$$\text{(a,d,e)} \quad \zeta = \{(p-1,q),(p,q-1),(p,q+1),(p+1,q)\}$$

$$\text{(b,c)} \quad \zeta = \{(p-1,q-1),(p-1,q+1),(p+1,q-1),(p+1,q+1)\}$$

Edge-sensing mechanism (ESM)

Essential to produce sharply looking images

- structural constraints impossed on the camera solution relate to the form of the ESM operator λ used to generate the edge-sensing weights

$$\lambda(z) \rightarrow \{w_{(i,j)}, (i,j) \in \zeta\}$$

- both structural and spatial characteristics are considered in the ESM construction

Concept

- ESM operator λ uses some form of inverse gradient of the samples in the CFA image

$$w_{(i,j)} = \frac{1}{1 + f(\Delta_{(i,j)})}$$

- large image gradients usually indicate that the corresponding vectors are located across edges (penalized through small weights)

Edge-sensing mechanism (ESM)

Conventional designs:

- **operate on large (5x5,7x7) neighbourhood**
- **specialization on a particular CFA (e.g. Bayer CFA):**

for shape mask $\zeta = \{(p-1,q),(p,q-1),(p,q+1),(p+1,q)\}$

$$w_{(p-1,q)} = 1/(1+|z_{(p-2,q)} - z_{(p,q)}| + |z_{(p-1,q)} - z_{(p+1,q)}|)$$
$$w_{(p,q-1)} = 1/(1+|z_{(p,q-2)} - z_{(p,q)}| + |z_{(p,q-1)} - z_{(p,q+1)}|)$$
$$w_{(p,q+1)} = 1/(1+|z_{(p,q+2)} - z_{(p,q)}| + |z_{(p,q+1)} - z_{(p,q-1)}|)$$
$$w_{(p+1,q)} = 1/(1+|z_{(p+2,q)} - z_{(p,q)}| + |z_{(p+1,q)} - z_{(p-1,q)}|)$$

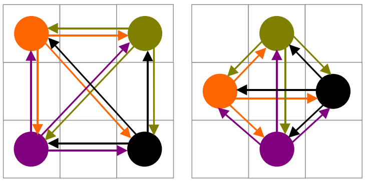
for shape mask $\zeta = \{(p-1,q-1),(p-1,q+1),(p+1,q-1),(p+1,q+1)\}$

$$w_{(p-1,q-1)} = 1/(1+|z_{(p-2,q-2)} - z_{(p,q)}| + |z_{(p-1,q-1)} - z_{(p+1,q+1)}|)$$
$$w_{(p-1,q+1)} = 1/(1+|z_{(p-2,q+2)} - z_{(p,q)}| + |z_{(p-1,q+1)} - z_{(p+1,q-1)}|)$$
$$w_{(p+1,q-1)} = 1/(1+|z_{(p+2,q-2)} - z_{(p,q)}| + |z_{(p+1,q-1)} - z_{(p-1,q+1)}|)$$
$$w_{(p+1,q+1)} = 1/(1+|z_{(p+2,q+2)} - z_{(p,q)}| + |z_{(p+1,q+1)} - z_{(p-1,q-1)}|)$$

Edge-sensing mechanism (ESM)

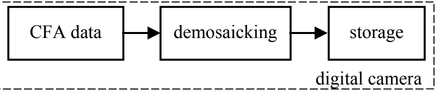
Cost-effective, universal design

- **operates within the shape mask ζ**
- **aggregation concept defined here over the four-neighborhoods only**
- **desing suitable for any existing CFA**

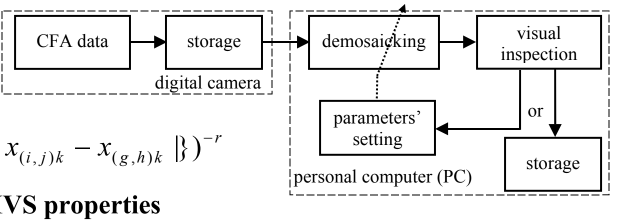


Fully automated solution

$$w_{(i,j)} = 1/(1+\sum_{(g,h)\in\zeta} |x_{(i,j)k} - x_{(g,h)k}|)$$



End-user control based solution



$$w_{(i,j)} = \beta(1+\exp\{\sum_{(g,h)\in\zeta} |x_{(i,j)k} - x_{(g,h)k}|\})^{-r}$$

- **matches better the HVS properties**

Spectral model (SM)

- **considers spectral & spatial characteristics of neighboring color pixels**

pixel occupying location to be interpolated $\mathbf{x}_{(p,q)} = [x_{(p,q)1}, x_{(p,q)2}, x_{(p,q)3}]$

pixel occupying neighboring location $\mathbf{x}_{(i,j)} = [x_{(i,j)1}, x_{(i,j)2}, x_{(i,j)3}]$

Modelling assumption in the existing SMs:

- **color ratio model (uniform hue modelling assumption)**

$x_{(p,q)k} / x_{(i,j)k} = x_{(p,q)2} / x_{(i,j)2}$; $k = 1$ or $k = 3$

- **normalized color ratio model (hue constancy is enforced in both in edge transitions and uniform image areas)**

$$(x_{(p,q)k} + \gamma)/(x_{(i,j)k} + \gamma) = (x_{(p,q)2} + \gamma)/(x_{(i,j)2} + \gamma)$$

- **color difference model (constrained component-wise magnitude difference)**

$$x_{(p,q)k} - x_{(i,j)k} = x_{(p,q)2} - x_{(i,j)2}$$

Vector SM

Modelling assumption

- **two neighboring vectors should have identical color chromaticity properties (directional characteristics)**
- **two spatially neighboring vectors should be collinear in the RGB (vector) color space**

Computational approach

$$\mathbf{x}_{(p,q)} . \mathbf{x}_{(i,j)} = \|\mathbf{x}_{(p,q)}\| \|\mathbf{x}_{(i,j)}\| \cos\left(\left\langle \mathbf{x}_{(p,q)}, \mathbf{x}_{(i,j)} \right\rangle\right)$$

$$\left\langle \mathbf{x}_{(p,q)}, \mathbf{x}_{(i,j)} \right\rangle = 0 \Leftrightarrow \frac{\sum_{k=1}^{3} x_{(p,q)k} x_{(i,j)k}}{\sqrt{\sum_{k=1}^{3} x_{(p,q)k}^2} \sqrt{\sum_{k=1}^{3} x_{(i,j)k}^2}} = 1$$

- **any color component can be determined from the expression above by solving the quadratic equation expression** $ay^2 + by + c = 0$
- ***y* denotes the component to be determined, e.g.** $y = x_{(p,q)2}$

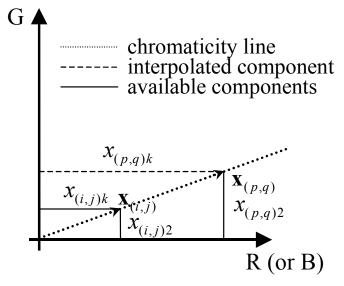
Vector SM

Unique quadratic equation solution

$$y = y_1 = y_2 = \frac{-b}{2a}$$ **due to zero discriminant** $$b^2 - 4ac = 0$$

Geometric interpretation

• **from two-component vector expression**



for G component

$$x_{(p,q)2} = \frac{x_{(p,q)k} x_{(i,j)2}}{x_{(i,j)k}}$$

for R or B component

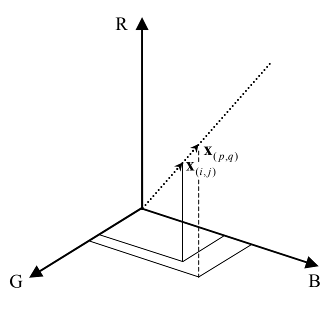
$$x_{(p,q)k} = \frac{x_{(p,q)2} x_{(i,j)k}}{x_{(i,j)2}}$$

Vector SM

Geometric interpretation

• **from three-component vector expression**

$$y = y_1 = y_2 = \frac{-b}{2a}$$



for G component

$$x_{(p,q)2} = \frac{x_{(p,q)1} x_{(i,j)1} x_{(i,j)2} + x_{(p,q)3} x_{(i,j)2} x_{(i,j)3}}{x_{(i,j)1}^2 + x_{(i,j)3}^2}$$

for R component

$$x_{(p,q)1} = \frac{x_{(p,q)2} x_{(i,j)1} x_{(i,j)2} + x_{(p,q)3} x_{(i,j)1} x_{(i,j)3}}{x_{(i,j)2}^2 + x_{(i,j)3}^2}$$

for B component

$$x_{(p,q)3} = \frac{x_{(p,q)1} x_{(i,j)1} x_{(i,j)3} + x_{(p,q)2} x_{(i,j)2} x_{(i,j)3}}{x_{(i,j)1}^2 + x_{(i,j)2}^2}$$

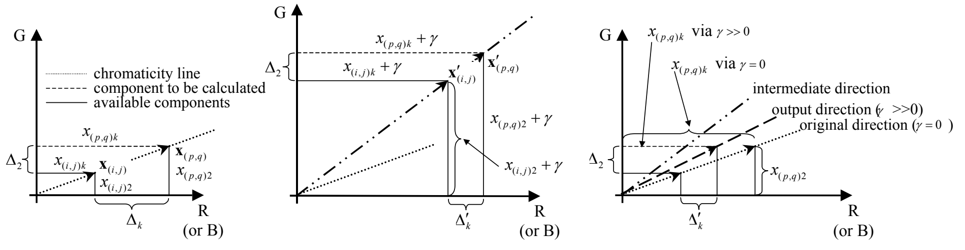
Generalized vector SM

Linear shifting of the input vectors

- **modifies their directional characteristics and normalizes their component-wise magnitude differences**

$$\left\langle \mathbf{x}_{(p,q)} + \gamma \mathbf{I}, \mathbf{x}_{(i,j)} + \gamma \mathbf{I} \right\rangle = 0 \quad \Leftrightarrow \quad \frac{[\mathbf{x}_{(p,q)} + \gamma \mathbf{I}].[\mathbf{x}_{(i,j)} + \gamma \mathbf{I}]}{\left\| \mathbf{x}_{(p,q)} + \gamma \mathbf{I} \right\| \left\| \mathbf{x}_{(i,j)} + \gamma \mathbf{I} \right\|} = 1$$

Geometric interpretation of 2-D case



Generalized vector SM

Features

- **universal solution: easy to implement**
- **tunes both directional & magnitude characteristics**
- **generalizes all previous spectral models:**

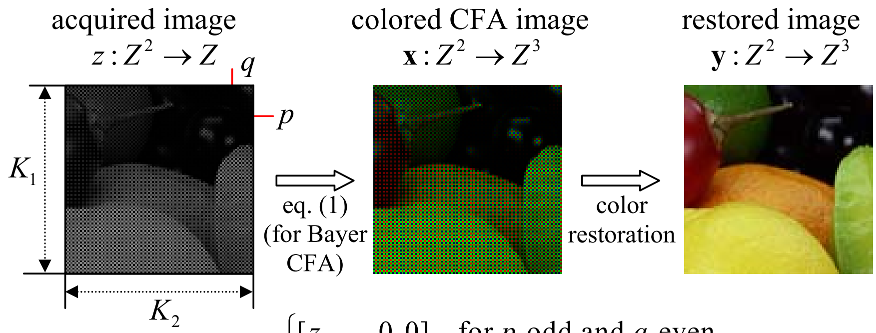
"non-shifted" vector model	($\gamma = 0$, three-component expression)
normalized color ratio model	(two-component expression)
color ratio model	($\gamma = 0$, two-component expression)
color difference model	($\gamma \to \infty$, two-component expression)

Vector SM based data-adaptive estimator

$$x_{(p,q)k} = \sum_{(i,j)\in\zeta} \{ w'_{(i,j)} x^{(i,j)}_{(p,q)k} \} \qquad x^{(i,j)}_{(p,q)k} = y - \gamma$$

Demosaicking (spectral interpolation)

From gray-scale input image to full- color output image

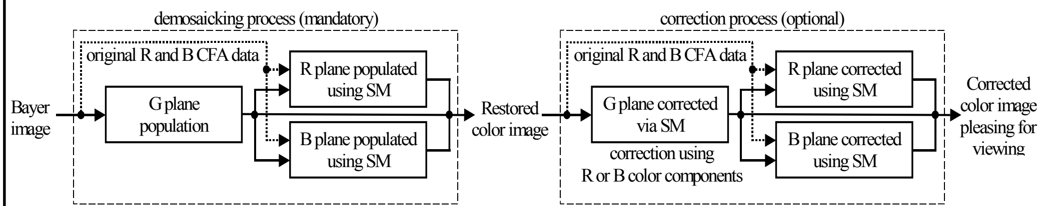


$$\mathbf{x}_{(p,q)} = \begin{cases} [z_{(p,q)}, 0, 0] & \text{for } p \text{ odd and } q \text{ even,} \\ [0, 0, z_{(p,q)}] & \text{for } p \text{ even and } q \text{ odd,} \\ [0, z_{(p,q)}, 0] & \text{otherwise} \end{cases} \qquad (1)$$

Demosaicking (spectral interpolation)

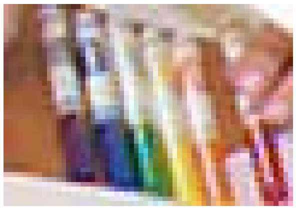
Color image: only with demosaicking

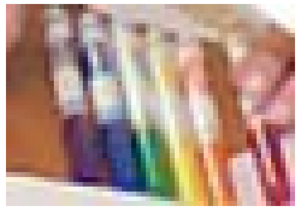
- **integral processing step in the pipeline**
- **should be supported by image post processing (correction)**



- **demosaicking vs. demosaicked image post processing: two fundamentaly different processing steps; they utilize similar, if not identical, signal processing concepts.**
- **post processing of demosaicked images: novel application**

SM and the ESM vs. color reconstruction quality

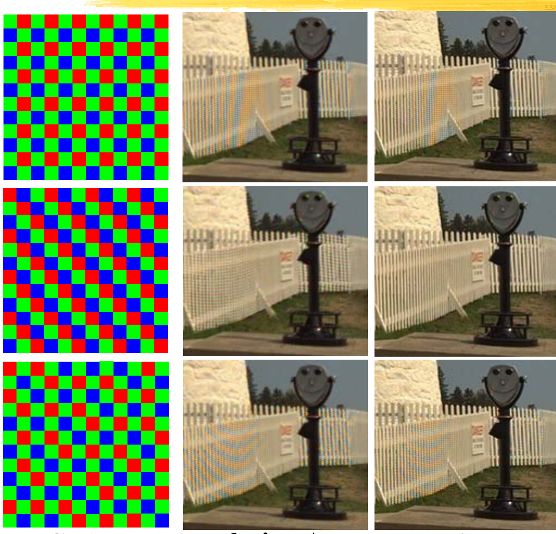

without SM and ESM


omitted SM, used ESM


omitted ESM, used SM


both SM and ESM used

CFA selection vs. demosaicking


CFA solution A solution B

Impact on image quality:

- quality significantly varies depending on both the CFA and the input image content

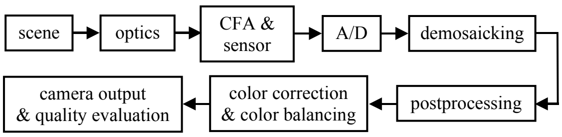
Impact on computational complexity:

- increased complexity for pseudo-random and random CFAs
- Bayer CFA offers one of the simplest color reconstruction

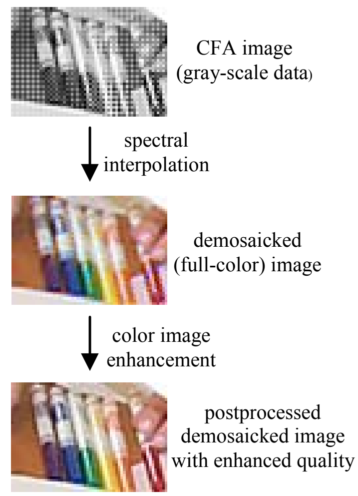
Demosaicked image post processing

Full-color image enhancement

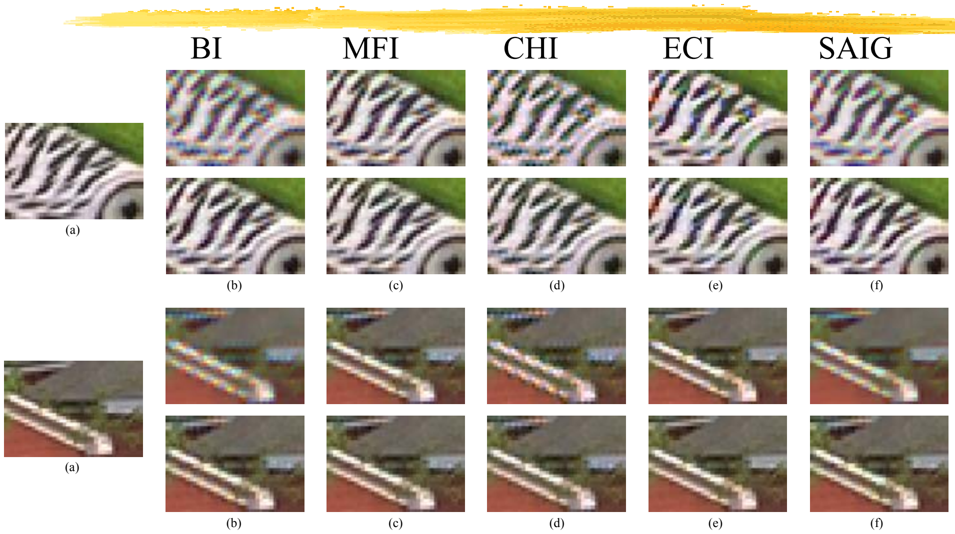
- postprocessing the demosaicked image is an optional step
- implemented mainly in software and activated by the end-user



- localizes and eliminates false colors created during demosaicking
- improves both the color appearance and the sharpness of the demosaicked image
- unlike demosaicking, postprocessing can be applied iteratively until certain quality criteria are met



Demosaicked image post processing



- demosaicked images (top rows), postprocessed images (bottom rows)

Digital zooming in imaging devices

Motivation

• technological advances -> miniaturization of single-sensor cameras

• pocket devices, mobile phones and PDAs -> low optical capabilities and computational resources

• to improve functionality and quality of output -> increase the spatial resolution of the camera output

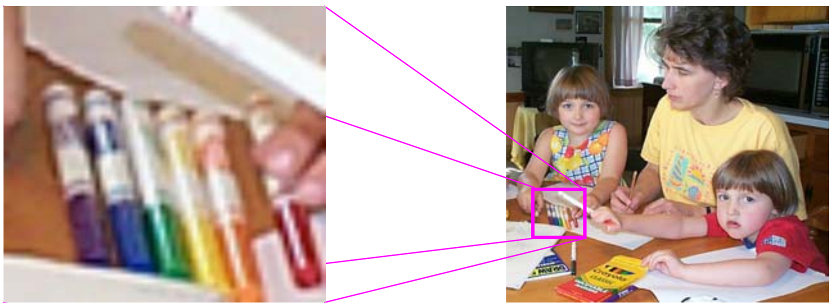


Image zooming (spatial interpolation)

Zooming in the RGB domain

• conventionally used

• slower - more samples to process

• amplification of the imperfections introduced during demosaicking

CFA image data → demosaicking → color image zooming → zoomed image

Zooming in the CFA domain

• novel approach

• operating on noise-free samples

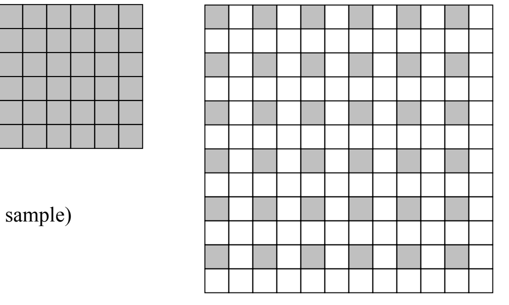
• spectral interpolation follows spatial interpolation

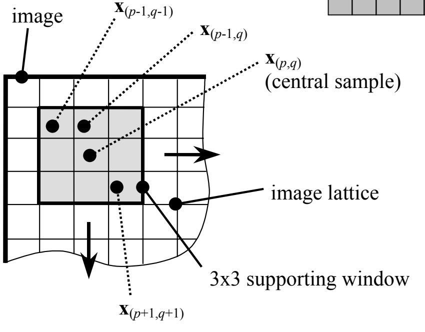
CFA image data → CFA image zooming → demosaicking → zoomed image

Demosaicked (full-color) image zooming

Zooming in the RGB domain

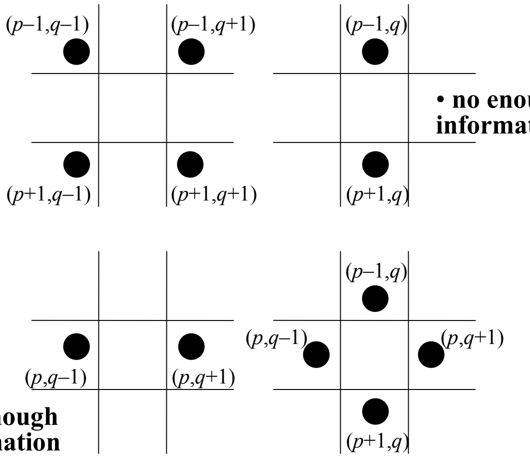
• **conventionally used**





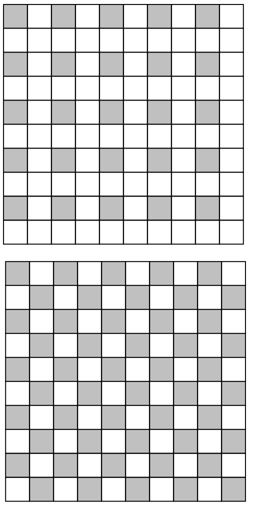
Demosaicked (full-color) image zooming

Pixel arrangements observed during processing



• **no enough information**

• **no enough information**



Demosaicked (full-color) image zooming

Zooming methods

- adaptive vs. non-adaptive
- component-wise vs. vector

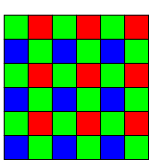
original | component-wise median | vector median

CFA image zooming

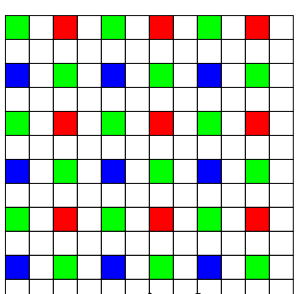
Filling CFA components

- **conventional approach destroys the underlying CFA structure**
- **specially designed "filling operations"**

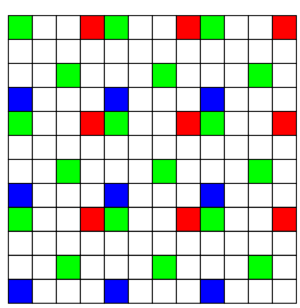
$$\left.\begin{array}{l}\mathbf{x}_{(2p-1,2q)}\\ \mathbf{x}_{(2p,2q-1)}\\ \mathbf{x}_{(2p-1,2q-1)}\end{array}\right\} = \mathbf{b}_{(p,q)} \quad \begin{array}{l}\text{for (odd } p, \text{even } q)\\ \text{for (even } p, \text{odd } q)\\ \text{otherwise}\end{array}$$



input Bayer CFA image



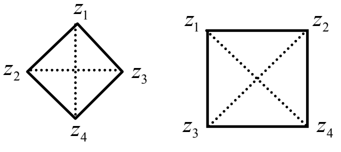
conventional



CFA based approach

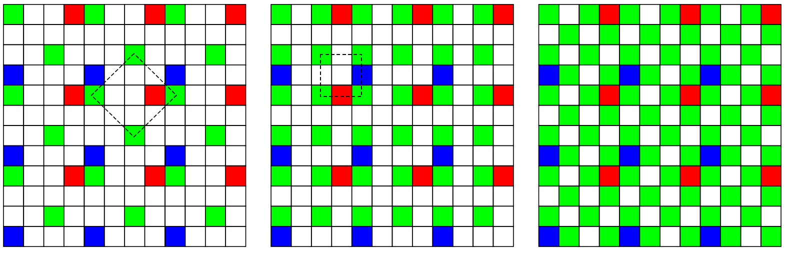
CFA image zooming

G interpolation step



interpolator $z_{(p,q)} = \sum_{j=1}^{4} w'_j z_j$

edge-sensing weight $w_i = \dfrac{1}{1 + \sum_{j=1}^{4} \left| z_i - z_j \right|}$



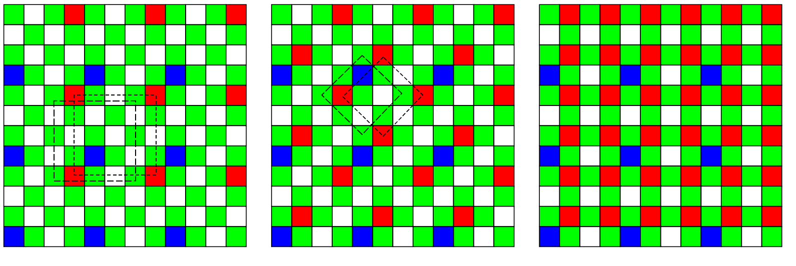
CFA image zooming

R interpolation steps

- **utilizes both spatial and spectral image characteristics**

$$z_{(p,q)} = z_{(p,q-1)} + \frac{\sum_{j=1}^{4} w_j \bar{z}_j}{\sum_{j=1}^{4} w_j} = z_{(p,q-1)} + \sum_{j=1}^{4} w'_j \bar{z}_j$$

spectral quantities are formed using spatially shifted samples $\bar{z}_i = R_i - G_i$



CFA image zooming

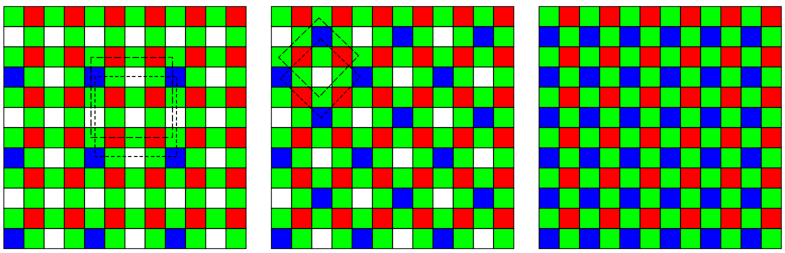
B interpolation steps

• **diagonal symmetry compared to R components**

$$z_{(p,q)} = z_{(p-1,q)} + \frac{\sum_{j=1}^{4} w_j \bar{z}_j}{\sum_{j=1}^{4} w_j} = z_{(p-1,q)} + \sum_{j=1}^{4} w'_j \bar{z}_j$$

spectral quantities are formed using spatially shifted samples

$$\bar{z}_i = B_i - G_i$$



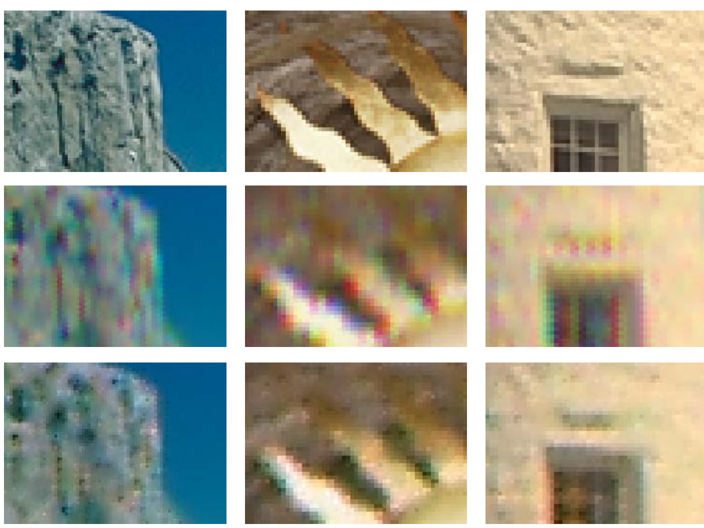
enlarged Bayer CFA image

Camera image zooming combined with demosaicking

• original images

• conventional (demosaicked) zooming

• CFA image zooming

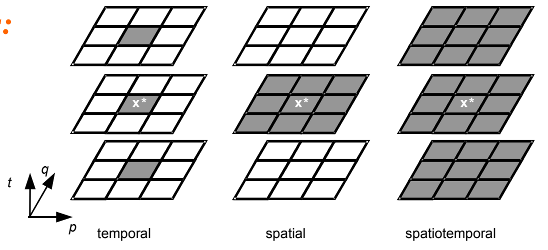


Video-demosaicking

Essential in single-sensor VIDEO cameras

- motion video or image sequences represent a 3-D image signal or a time sequence of 2-D images (frames)
- motion video usually exhibits significant correlation in both the spatial and temporal sense
- by omitting the essential temporal characteristics, spatial processing methods, which process separately the individual frames, produce an output image sequence with motion artifacts

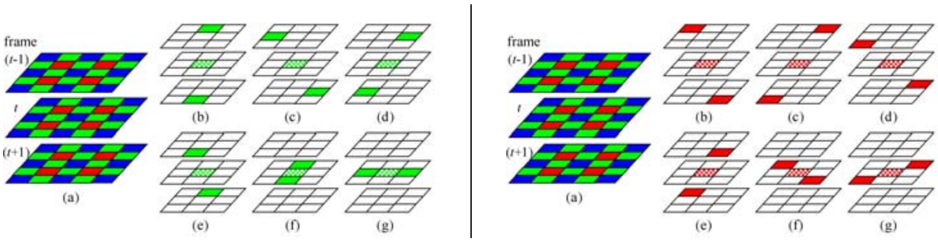
Processing windows:



Spatiotemporal video-demosaicking

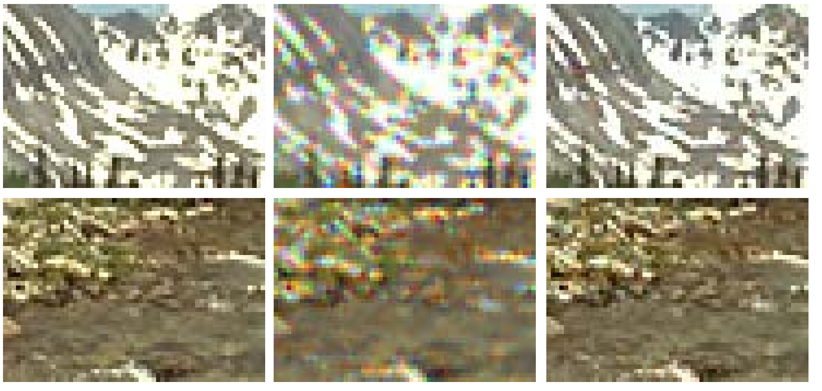
Fast video-demosaicking procedure

- usage in PDAs and mobile phone imaging applications
- utilization of multistage unidirectional spatiotemporal filtering concepts



- essential spectral quantities formed over the spatiotemporal neighborhood
- structural content followed by spatiotemporal edge-sensing weights
- color component to be outputted is obtained via weighted average operations defined over unidirectional demosaicked values

Video-demosaicking



original frames | restored using spatial BI demosaicking | restored using fast spatiotemporal demosaicking

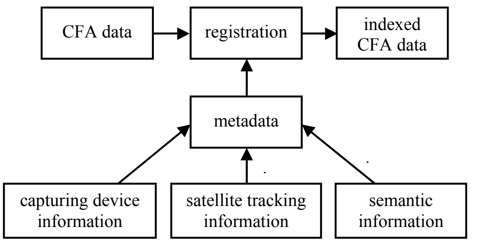
Video-demosaicking

original frames | restored using spatial BI demosaicking | restored using fast spatiotemporal demosaicking

Camera image indexing

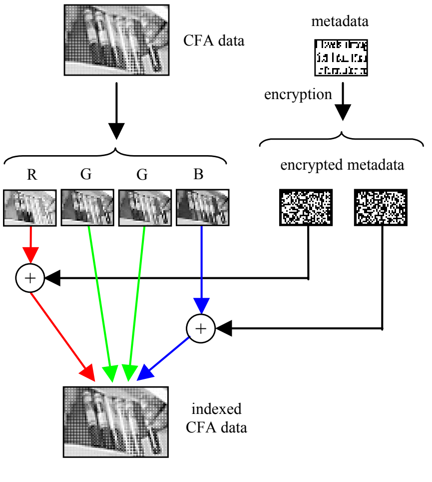
Digital rights management in digital cameras:

- captured images are directly indexed in the single sensor digital camera, mobile phone and pocket device
- indexing performed by embedding metadata information
- great importance to the end-users, database software programmers, and consumer electronics manufacturers

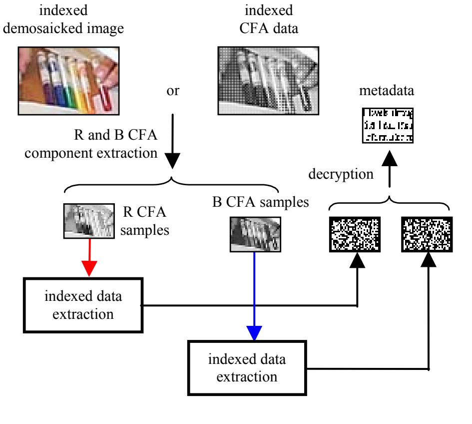


Camera image indexing

Embedding procedure

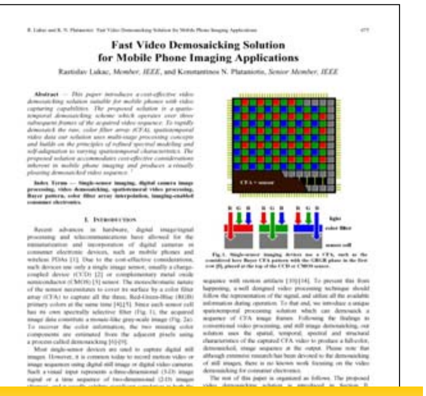


Extraction procedure



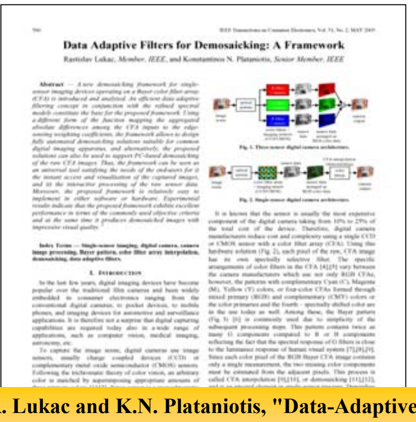
Where to learn more?

R. Lukac, B. Smolka, K. Martin, K.N. Plataniotis, and A.N. Venetsanopoulos, "Vector Filtering for Color Imaging," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 74-86, January 2005.

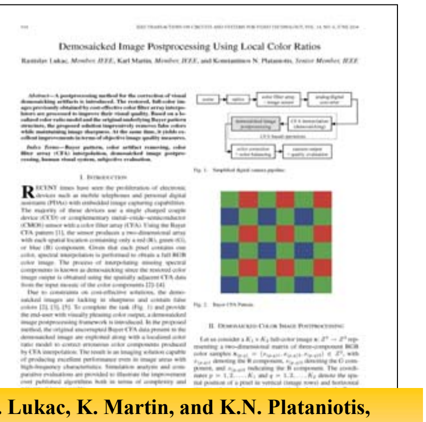


R. Lukac and K.N. Plataniotis, "Fast Video Demosaicking Solution for Mobile Phone Imaging Applications," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 2, pp. 675-681, May 2005.

Where to learn more?

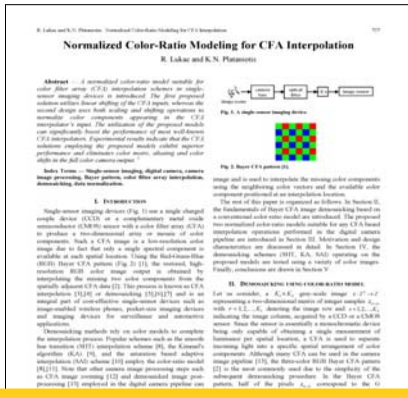


R. Lukac and K.N. Plataniotis, "Data-Adaptive Filters for Demosaicking: A Framework," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 2, pp. 560-570, May 2005.



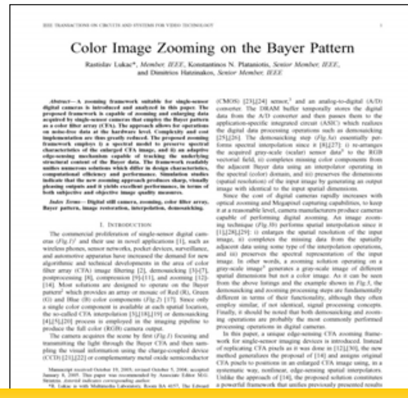
R. Lukac, K. Martin, and K.N. Plataniotis, "Demosaicked Image Postprocessing Using Local Color Ratios," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 6, pp. 914-920, June 2004.

Where to learn more?



Normalized Color-Ratio Modeling for CFA Interpolation

R. Lukac and K.N. Plataniotis

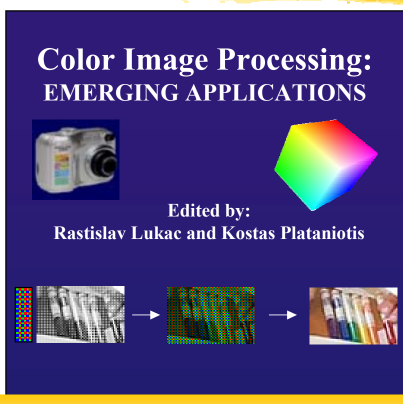


Color Image Zooming on the Bayer Pattern

R. Lukac and K.N. Plataniotis, "Normalized Color-Ratio Modelling for CFA Interpolation," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 2, pp. 737-745, May 2004.

R. Lukac, K.N. Plataniotis, and D. Hatzinakos, "Color Image Zooming on the Bayer Pattern," *IEEE Transactions on Circuits and Systems for Video Technology*, to appear, vol. 15, 2005.

Where to learn more?



R. Lukac and K.N. Plataniotis, "*Color Image Processing: Emerging Applications*," *CRC Press*, spring 2006.

www.dsp.utoronto.ca/~lukacr/index.php?page=research3