Demosaicing of Color Images Using Steerable Wavelets

Yacov Hel-Or, Daniel Keren
HP Laboratories Israel
HPL-2002-206 (R.1)
August 9th , 2002*

E-mail: toky@hp.technion.ac.il

In some types of digital color cameras only a single value is provided for each pixel: either, red, green, or blue. The reconstruction of the three color values for each pixel is known as the "demosaicing" problem. This report suggests a solution to this problem based on the steerable wavelet decomposition. The reconstructed images present a faithful interpolation of missing data while preserving edges and textures.

Demosaicing of Color Images Using Steerable Wavelets

Yacov Hel-Or and Daniel Keren
Hewlett-Packard Labs Israel *

Abstract

In some types of digital color cameras only a single value is provided for each pixel: either, red, green, or blue. The reconstruction of the three color values for each pixel is known as the "demosaicing" problem. This report suggests a solution to this problem based on the steerable wavelet decomposition. The reconstructed images present a faithful interpolation of missing data while preserving edges and textures.

1   Introduction

A color image is typically represented by three bands each of which is a matrix of values representing the responses of an array of photo-sensors to the viewed scene. The three bands are often referred to as red (R), green (G), and blue (B) according to their spectral sensitivity. Thus, three numbers are given at each matrix location composing a *pixel* value. Some CCD cameras, however, provide only a single value for each pixel due to their inability to position three photo-sensors at the same location. In these cases, the captured image is composed of three bands whose pixel values are partially sampled. An example of a CCD array with a typical photo-sensor arrangement is given in Figure 1.

The *demosicing* problem deals with the reconstruction of a color image I from a partial sampling D of its pixel values. This inverse problem is, of course, under-determine since the solution space includes infinite images satisfying D. To obtain a single solution, a quality criterion should be defined over the solution space so that the best image according to that

---

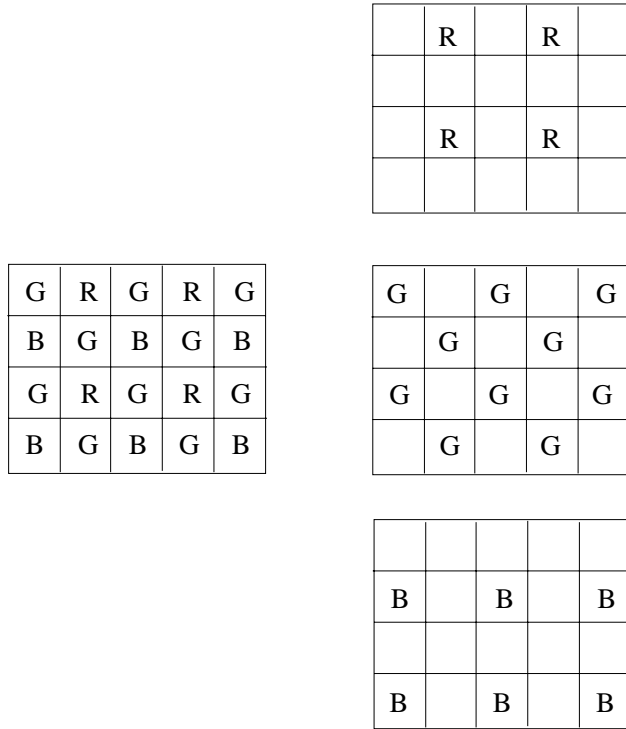*Address: HP Labs Israel, Technion City, Haifa 32000, Israel. Email: toky@hp.technion.ac.il

	R		R	
	R		R	

G	R	G	R	G
B	G	B	G	B
G	R	G	R	G
B	G	B	G	B

G		G		G
	G		G	
G		G		G
	G		G	

B		B		B
B		B		B

Figure 1: A typical arrangement of photo-sensors in a CCD array (left). This arrangement provides three color bands whose pixel values are partially sampled (right). Note that in this example, the number of green samples is twice that of the red and the blue. This corresponds to the varying spatial resolution of the human visual system at different spectral wavelength.

criterion will be chosen. Assuming $P(I)$ specifies the probability that an image I is the acquired image, the *maximum a posteriori* (MAP) criterion defines the desired solution as the image $\hat{I}$ that maximizes the *a posteriori* conditional probability:

$$\hat{I} = \arg\max_{I} P(I|D) \tag{1}$$

In general it is difficult to evaluate $P(I|D)$ and the Bayes rule is applied to this end:

$$P(I|D) = \frac{P(D|I)P(I)}{P(D)} \tag{2}$$

Since D is already given, $P(D)$ is constant and can be disregarded. As a result, estimating the *a posteriori* probability $P(I|D)$ requires the evaluation of $P(D|I)$ and $P(I)$. The first term is typically easy to define and depends on the noise characteristics of the measured samples. On the other hand, modeling the *a priori* probability $P(I)$ is a complicated task and can be critical for any reconstruction approach.

Previous studies dealing with the demosaic problem suggested several alternatives to specify an image prior $P(I)$. Brainard and Sherman [BS95, Bra94] use the spectral characteristics of natural images and the correlation between color bands to define a prior over the image population. With their approach the prior function is image independent and is calculated once and for all natural images. Keren [Ker96] suggested the use of an adaptive prior that is calculated for each specific image and changes over the location in the image based on the partial information available. Both approaches assume the smoothness of natural images in the spatial domain. This smoothness assumption is problematic in textured areas and near edges or lines where the smoothness assumption is incorrect. Applying the smoothness prior to these areas will result in blurred edges and a loss of high frequency details.

This report suggests a new approach which exploits the correlation between neighboring information in the "steerable wavelet" representation. The advantage of working in the wavelet domain is that the smoothness assumption is applied in the appropriate scale and in the appropriate orientation. The demosaic results using this approach provides a faithful interpolation of missing samples while preserving edges and textures in the resulting image.

The report is organized as follows: First, the basic problem and the suggested solution are presented for a simpler case; the reconstruction of a gray-scale images from its partial sampling. Next, the steerable wavelet decomposition is presented. Finally, the reconstruction of color images using the steerable wavelet decomposition is explained. It should be noted that though the current paper is written in the the context of demosaicing solutions, the suggested approach is also appropriate for a wide variety of reconstruction problems that can be dealt with using the same framework.

2   Modeling Gray-Scale Natural Images

Many psychophysical and computational studies show that natural images have special characteristics forming a non uniform distribution over the image space (see e.g. [Ker87, OF96]). For example, it is widely assumed that the power spectra of natural images tend to be inversely proportional to their spatial frequencies [Fie93]. Additionally, it is shown that natural images are spatially and spectrally correlated within neighborhood areas [OF96]. In

color images, the correlation is even more restrictive due to the fact that there is a strong correlation (or anti-correlation) between color bands [Wan95, HK84].

2.1 Isotropic smoothness Model

Previous approaches for the reconstruction of a single-band image used the Markov Random Field (MRF) model for introducing the prior probability $P(I)$ [GG84, MMP87]. With this model, natural images are assumed to be isotropically "smooth" with probability conforming to the Gibbs distribution:

$$P(I) = \frac{1}{Z} e^{-\Psi(I)}$$

where Z is a normalization constant and $\Psi(I)$ is a smoothness measure of the image I. Popular smoothness measures are the first order (membrane) and the second order (thin plate) energy forms:

$$\Psi_1(I) = \int_\Omega \|\nabla I\|^2 d\Omega$$

$$\Psi_2(I) = \int_\Omega \|\nabla^2 I\|^2 d\Omega$$

where the integral is performed over the image support Ω.

Assuming the sampled values are $D(x_i, y_i)$ and the measurement noise is a white zero mean Gaussian process with variance σ^2, the conditional probability $P(D|I)$ then follows the Gaussian distribution:

$$P(D|I) = \frac{1}{Z'} \exp\{-\frac{\sum_i (D(x_i, y_i) - I(x_i, y_i)^2)}{2\sigma^2}\}$$

Using, for example, the second order MRF model, and plugging $P(D|I)$ and $P(I)$ above into Equation 1 we obtain:

$$\hat{I} = \arg\max_I \left[\frac{1}{Z''} \exp\left\{ -\frac{\sum_i (D(x_i, y_i) - I(x_i, y_i)^2)}{2\sigma^2} - \int_\Omega \|\nabla^2 I\|^2 d\Omega \right\} \right]$$

It is easy to verify that maximizing the above expression is equivalent to minimizing:

$$C(I|D) = \frac{\sum_i (D(x_i, y_i) - I(x_i, y_i)^2)}{2\sigma^2} + \lambda \int_\Omega \|\nabla^2 I\|^2 d\Omega \qquad (3)$$

This *cost function* consists of two terms: the first term penalizes pixel values not conforming with the sampled values, and the second term penalizes for a non smooth solution. The variable λ is a tuning parameter that determines the relative weight of the smoothing term in the total cost.

Minimizing Equation 3 yields a solution that compromises between the sampling values and the smoothness term. Due to the prior assumption that natural images are isotropically "smooth", areas in the image that are indeed "smooth" will be reliably reconstructed. However, edges, lines, and textured areas in the original image that do not conform with the prior model will be blurred. Moreover, if the sampling values are assumed to be error free (i.e. noise variance is zero), oscillating artifacts will appear near edges. This is demonstrated in Figure 2. The upper figure shows partial sampling of the green band as taken by a digital camera Kodak DCS-200. The lower figure shows the reconstructed image using an iterative scheme that minimizes Equation 3 where the sampled values are assumed to be error free. The oscillating artifacts near edges and lines are demonstrated (see, for example, the net pattern in between the rungs in the chair back).
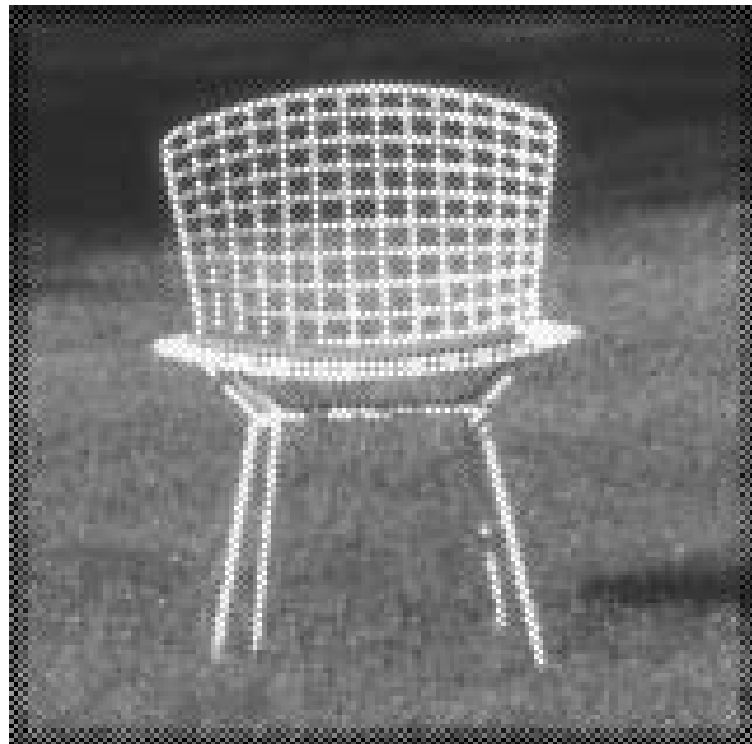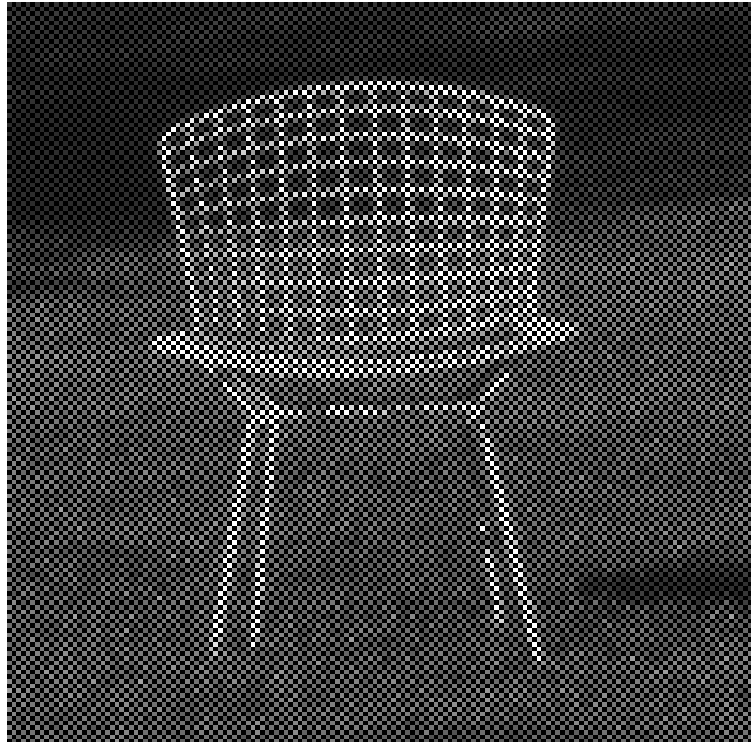
Figure 2: Upper figure: The green band of a sampled image taken by a digital camera (Kodak DCS-200). Lower figure: Reconstructed figure using the isotropic smoothness assumption.

2.2   Directional Smoothness Model

As mentioned in the previous section, the main problem with the isotropic smoothness model is the interpolation near edges and in textured areas where the smoothness assumption is incorrect. For this end, several alternatives were suggested, all of them model a typical image as a piecewise constant or a piecewise linear function. Anisotropic diffusion [Pm90], robust statistics [MMRK91], and regularization with line processes [GG84], were developed to treat image interpolation and reconstruction with discontinuities. In recent studies [BZ87, BR96, BSmH97] all the above approaches were shown to be theoretically equivalent.

This work is similar in spirit to these approaches, namely, interpolation is performed along edges and not across edges. Previous methods are based on the localization of edges using gradient operations and preventing the interpolation from crossing these edges. This leads to two main difficulties: The gradient is a too local feature and thus produces many spurious edges especially when the data is sparsely sampled as in our case. Additionally, these approaches operate on a single scale of the image. Since edges can appear in different scales, only part of the edges are correctly treated. In this work the interpolation is based on the *orientation* of the underlying edges which is a more global feature and therefore less sensitive to missing data and noise. Furthermore, the interpolation is performed in a multi-scale manner and thus deals with each edge in its appropriate scale.

In this work a more intuitive prior assumption is given to natural images. Natural images are considered as $2D$ functions with sharp edges at different *scales* and different *orientations*. This prior is in accord with statistical [OF96], neuro-physiological [SGVE89], and psychophysical [BA91] studies where natural images are said to contain localized, oriented, and bandpass structures.

In order to efficiently obtain the particular structure of a given image with respect to the above properties a multi-scale and multi-orientation representation should be applied to the image. The multi-scale representation can be implemented using the *Laplacian pyramid* decomposition [BA83]. The Laplacian pyramid decomposes an image I into a set of band-

pass components:

$$I = \sum_s B_s$$

where each band B_s represents a particular scale in the Laplacian pyramid. The multi-orientation representation can be applied using *steerable filters* decomposition where B_s is decomposed into a linear set of oriented bands:

$$B_s = \sum_d B_{s,d}$$

where the sum is performed over the orientation bands $d = 0..m - 1$, and the preferred direction of each band is $\theta_d = \frac{\pi d}{m}$. The entire decomposition with respect to scale and orientation is referred to as the *steerable wavelet* decomposition [FA90]. The main advantage of the steerable wavelet decomposition is that the image response can be recovered for *any* direction. Thus, if $B_{s,d}$ represents the image response along the preferred direction θ_d then the image response along an arbitrary direction θ can be calculated by a weighted sum of $\{B_{s,d}\}$:

$$B_{s,\theta} = \sum_{k=0}^{m-1} h_k(\theta) B_{s,d}$$

where h_k are known functions that depend on the the angle θ. Typically, it is common to assume that high value of a particular coefficient $B_{s,d}(p)$ suggests for an oriented structure in the image, located at p with direction θ_d and scale s. However, this is not always true since the transform coefficient is phase dependent. In order to nullify the phase factor it is necessary to apply the *Hilbert transform*[1] of the steerable wavelets as well [Bra78]. A pair of filters which are Hilbert transforms of each other form a *quadrature pair*. Thus, for each scale and orientation we have a pair of bands $B_{s,d}$ and $\hat{B}_{s,d}$ from which an oriented structure can be deduced. The squared sum of two coefficients forming a quadrature pair represents the (phase independent) energy of the signal at the particular frequencies represented by the quadrature filters. In our case, a high energy response in:

$$E_{s,d}(p) = B_{s,d}^2(p) + \hat{B}_{s,d}^2(p)$$

indicates an oriented structure in the image, located at p with direction θ_d and scale s. If the Hilbert bands are steerable as well (and we design them to be such) then it is possible

---

[1] Hilbert transform of a filter has the same frequency response with a phase shift of $90°$.

to "steer" the energy response in a similar manner for any orientation θ:

$$E_{s,\theta}(p) = B_{s,\theta}^2(p) + \hat{B}_{s,\theta}^2(p)$$

In this manner, the *directional energy* which is the energy response at a particular preferred direction and in a particular scale can be calculated.

An example of steerable wavelet decomposition applied to a naive interpolation (bilinear) of the chair image is presented in Figure 3. In this decomposition we have chosen three scales and four orientations. The wavelet bands $\{B_{s,d}\}$ along with their quadrature pairs are presented. Figure 4 shows the energy response of each band calculated by a squared sum of its quadrature pairs. Bright areas indicate an oriented structure at the particular direction and scale.

Under the assumption that natural images contain sharp edges at different scales and different orientations, it is very probable that the coefficient values in the Laplacian pyramid will have similar values along directions with high directional energy responses. I.e. if $E_{s,\theta}$ is high, and a neighboring pixel q is at direction θ from p, then the difference $|B_s(p) - B_s(q)|$ should be small. This assumption can be expressed as a prior probability for a Laplacian band as follows:

$$P(B_s) = \frac{1}{Z} e^{-\Psi_3(B_s)}$$

where Ψ_3 is a directional smoothness function measuring the extent of smoothness of the coefficient band B_s according to the directional energy $E_{s,\theta}$. For example, adopting the membrane energy form for smoothing, the directional smoothness measure will be:

$$\Psi_3(B_s) = \sum_p \sum_{q \in N(p)} g(E_{s,\vec{pq}}(p))(B_s(p) - B_s(q))^2$$

where $N(p)$ is the local neighborhood of p, and $\vec{pq}$ is the angle representing the direction of q from p. The function $g(z)$ is an increasing monotonic function adjusting the strength of this smoothness criterion according to the directional energy. Assuming the Laplacian coefficients are uncorrelated along scales we have that the joint probability of the entire set of Laplacian coefficients $B = \{B_s\}$ is a chain multiplication of priors:

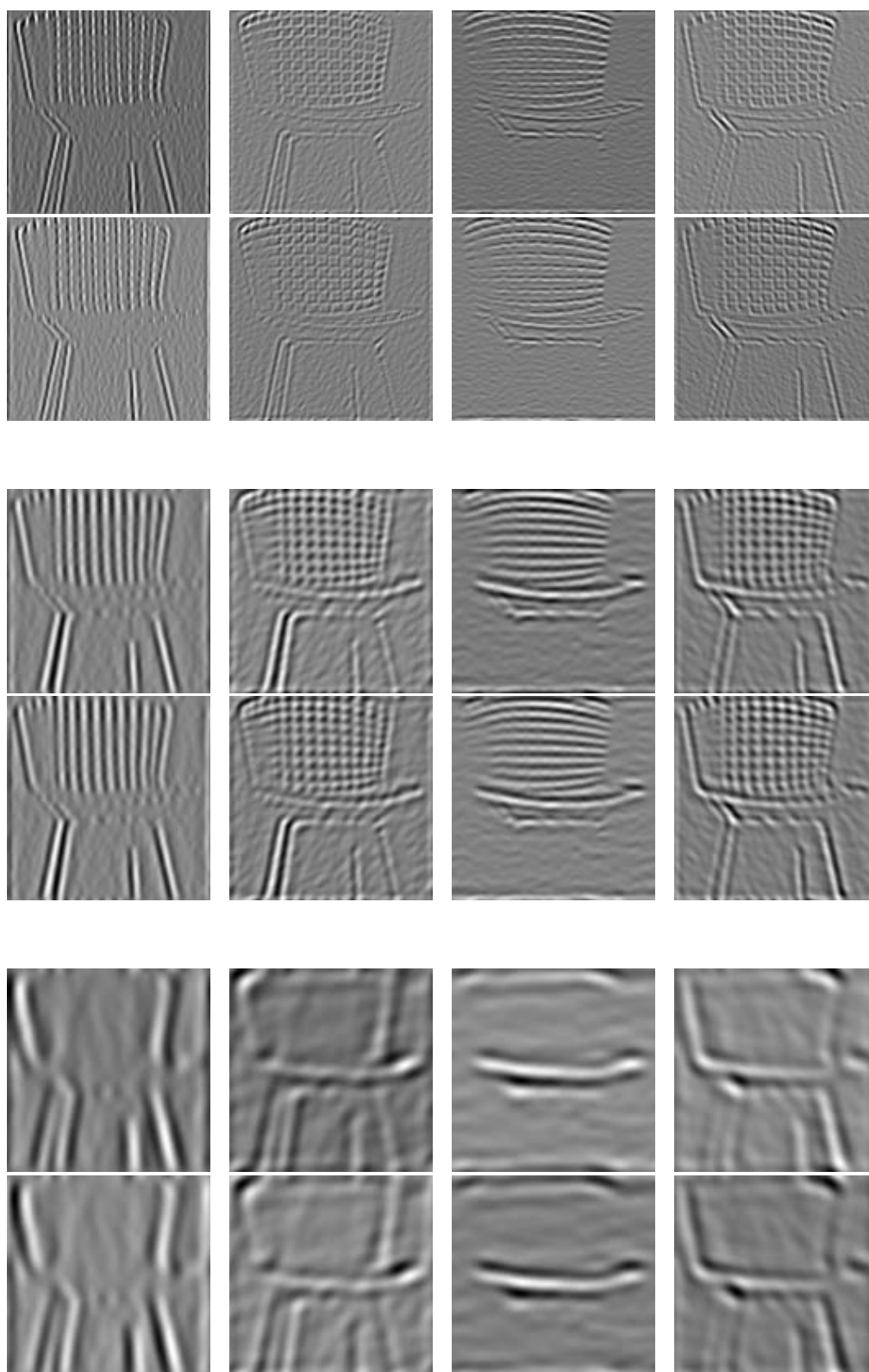$$P(B_1, B_2, \cdots) = P(B) = \prod_s P(B_s)$$

9

Figure 3: The Steerable Wavelets decomposition applied to a naive interpolation of the chair image. Three scales are presented (top to bottom: 1st, 2nd, and the 3rd octaves). At each scale eight steerable filters were applied; four directions were sampled where two steerable filters (forming a quadrature pair) were applied at each direction.
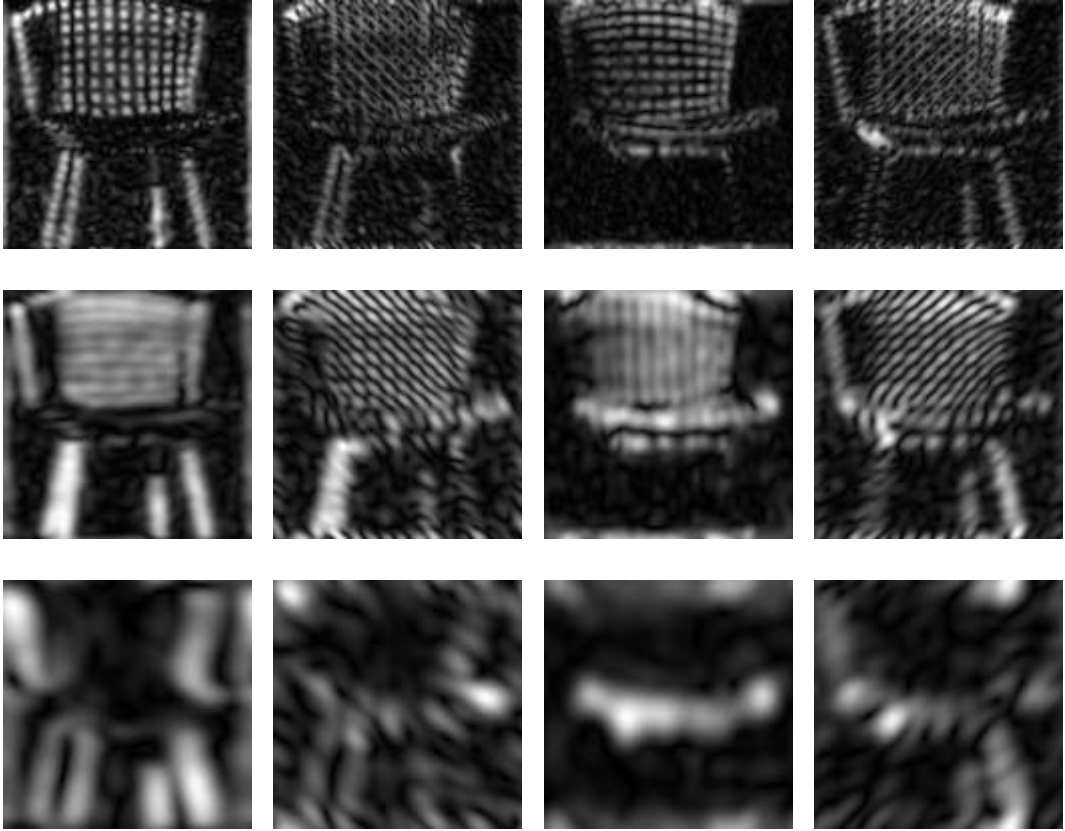
Figure 4: An energy response calculated by the squared sum of the quadrature values. Three scales (top to bottom) and four orientation (left to right) are presented. Bright areas indicate an oriented structure at the particular orientation and scale.

Since the decomposed image I defines uniquely its Laplacian coefficients (even though the Laplacian forms an over-complete representation) the prior distribution of B is proportional to the prior of natural images:

$$P(I) = \int P(I|B)P(B)dB = \int \delta(I - \sum_s B_s)P(B)dB \propto P(B^*) \quad \text{where} \quad I = \sum_s B_s^*$$

Plugging this into Equations 2 and 3 we find that minimizing the cost function $C(I|D)$ (Equation 3) can be performed alternatively by minimizing a corresponding term involving the Laplacian coefficients B:

$$C(I|D) \propto C(B|D) = \frac{\sum_i (D(x_i, y_i) - I(x_i, y_i)^2)}{2\sigma^2} + \lambda \sum_s \Psi_3(B_s^*) \tag{4}$$

The minimization of $C(I|D)$ can be performed iteratively using a gradient descent procedure. The iterations are performed in two steps minimizing the data term and the smoothness term

alternately. If B_s^t is the set of the Laplacian coefficients at iteration t, the following iterative scheme will minimize the smoothness term:

$$B_s^{t+1}(p) = B_s^t(p) + \frac{\beta}{Z} \sum_{q \in N(p)} g(E_{s,\vec{pq}}(p))(B_s^t(q) - B_s^t(p))$$

where β is a scalar parameter adjusting the convergence rate, and Z is a normalization constant $Z = \sum_{q \in N(p)} g(E_{s,\vec{pq}}(p))$. Note that the difference $B_s^t(q) - B_s^t(p)$ is weighted by $E_{s,\vec{pq}}$, thus, at each location p the Laplacian band B_s is directionally diffused along orientations with high directional energies. A similar iteration procedure is performed in order to minimize the data term. At step t the following iteration is performed:

$$I^{t+1}(p) = I^t(p) + \gamma(D(p) - I^t(p))$$

The scalar γ is similarly a convergence parameter depending on the error variance σ. If $D(p)$ is error free ($\sigma = 0$ as we assume in our case) γ is one and the iteration rule becomes: $I^{t+1}(p) = D(p)$. In this case, combining the two steps into one minimization procedure results in an algorithm as illustrated in Figure 5.

The effect of smoothing along the preferred orientations is demonstrated in Figures 6 and 7. In Figure 6 the diffusion was performed only along the vertical direction. It can be seen that vertical lines and edges were reliably interpolated while other edges remained unchanged. Figure 7 shows a diffusion along all directions where four directional bands were used. In this case edges and lines in all orientations were reliably interpolated.

12

Raw Image

↓

Naive interpolation

Laplacian Pyramid   Steerable Pyramid

for each scale
smooth along
prefered orientations

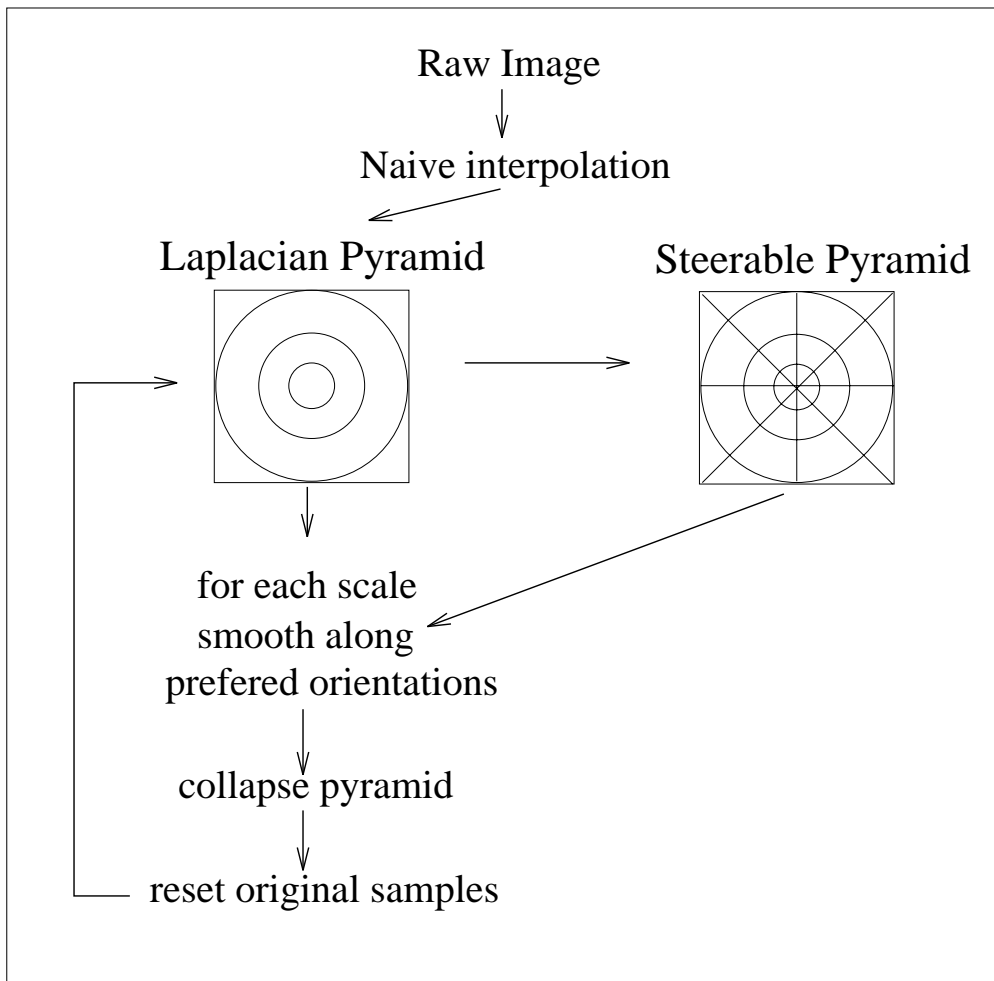collapse pyramid

reset original samples

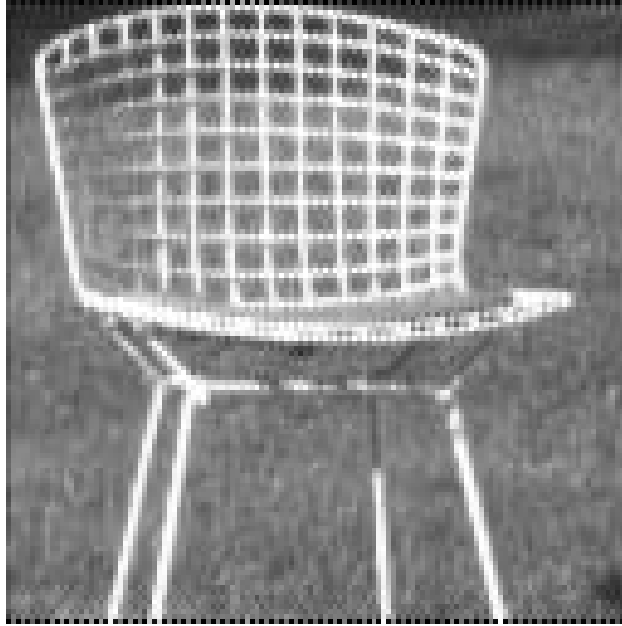Figure 5: Domosaicing algorithm for error free sampled data.

Figure 6: Directional interpolation of the green band taken by the Kodak DCS-200 camera. The interpolation was performed only for those bands whose preferred orientation were vertically oriented. The rest of the image remains unchanged.

Figure 7: Directional interpolation of the green band taken by the Kodak DCS-200 camera. The interpolation was perform for all directions.

3   Steerable Wavelets and Local Orientation

The wavelet transform has become a popular image representation in recent years. The transform gained its popularity due to several useful properties:

- Its basis functions are localized both in spatial and frequency domains.

- It decomposes an image into multi-scale and multi-orientation components which is a useful property in many processing tasks.

- It exploits the local correlations in natural images (signals), and thus its coefficients are (relatively) uncorrelated.

The wavelet decomposition is appropriate for our task since its coefficients capture multi-scale and multi-orientation structures in the image. However, the conventional Wavelet transform is inappropriate to this end due to the fact that it is not shift-invariant nor rotation-invariant. Figure 8 demonstrates this problem (taken from [SFAH92]).

Actually, every critically sampled transform will be variant to shift or rotation [SFAH92]. To overcome this problem it is essential to perform an over-complete transform (the number of coefficients is greater than the number of pixels). This is not a problem in our case since our aim is image analysis and not image coding.

The *steerable wavelet* decomposition [FA90] is a wavelet-like transform that is invariant to shift and rotation and thus appropriate for our task. It is an over-complete representation having the following properties:

- Its basis functions are shifted, scaled, and rotated versions of a "mother wavelet"; If $M(x, y)$ is a mother wavelet, then each basis kernel $W_{s,d}$ can be constructed by:

$$W_{s,d} = R_{\theta_d} M(sx, sy)$$

where (sx, sy) is a scaling by s, and R_{θ_d} represents a rotation by θ_d. Each wavelet band $B_{s,d}$ is a result of convolving the image I with the corresponding basis kernel:
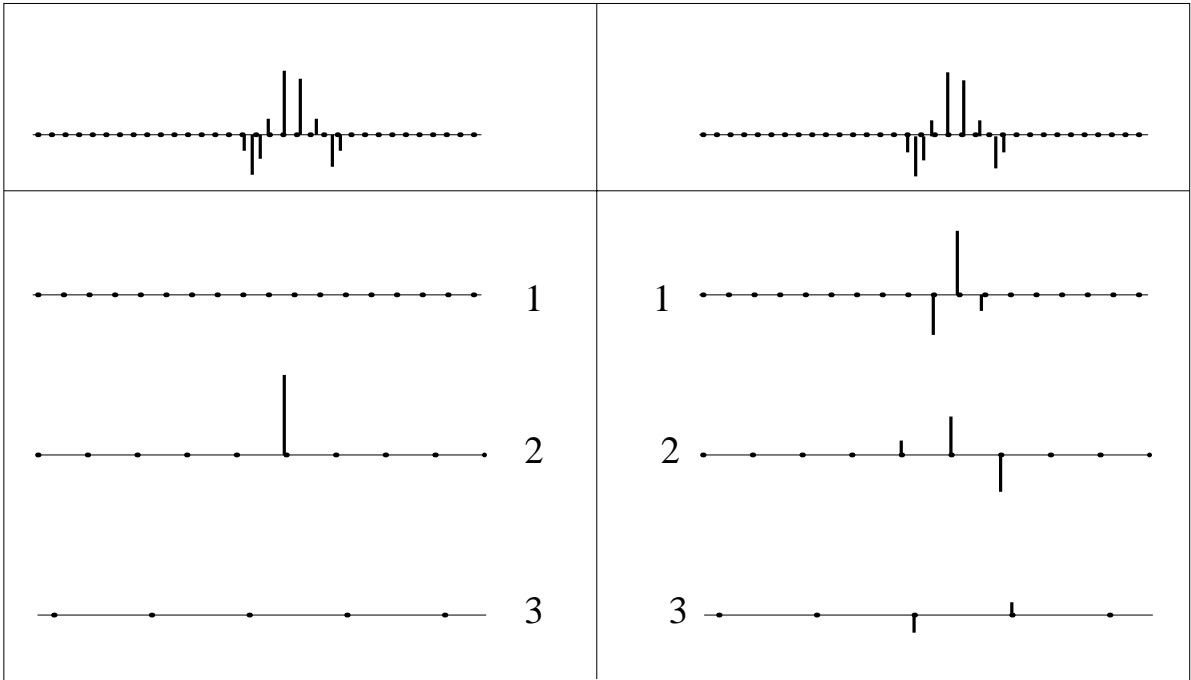
$$B_{s,d} = I * W_{s,d}$$

15

Figure 8: Assume the signal to be transformed (upper left) is composed of a particular wavelet basis function $w_k(x)$. All of the wavelet coefficients of this signal will have zero values except for the coefficient corresponding to the kth basis function. Now, if we shift the signal by one pixel (upper right), the coefficients of the new signal might be distributed in all the wavelet levels (right). In translation invariant transforms the coefficients of the translated signal are similarly translated. This example is taken from [Simoncelli et. al. 1992] .

- It is a self-invertible transform (the synthesis functions and the analysis functions are identical).

- It is shift and rotation invariant. Therefore, its coefficients form smooth structures and graceful changes in the wavelet domain.

- It has a flat response (the sum of squares of the basis functions is constant for all (x, y)).

- Its basis functions are similar to the the receptive fields of the simple cells in the visual cortex.

- Most important: it is possible to recover the filter response in *any* direction. I.e. if the

16

bands $B_{s,d}$ ($d = 0..m - 1$) are the signal responses in preferred directions θ_d then

$$\forall \theta \in [0 \cdots 2\pi] \quad B_{s,\theta} = \sum_{d=0}^{m-1} h_d(\theta) B_{s,d}$$

where $\{h_d\}_{d=0}^{m-1}$ are known interpolation functions. This is the "steerability" property of these filters.

The last property is very useful is our case. The "steerability" property allows us to calculate the directional energy of an image at any chosen direction by "steering" the directional energy. In [FA91] Freeman and Adelson explain how to calculate the directional energy from a set of steerable filters and their Hilbert transforms (forming quadrature pairs). The directional energy is phase-independent due to the phase asymmetry of the quadrature filter pair [Bra78]. Freeman and Adelson [FA91] suggested a set of steerable quadrature pairs based on derivatives of a Gaussian. They provide filter taps with the appropriate interpolation functions $\{h_d\}$.

An alternative approach for calculating the direction with the highest energy response, denoted by *local orientation*, is suggested by Haglund et. al. [HKG93]. The local orientation $\mathbf{d}_s(p)$ is a vector associated with the position p and scale s denoting the dominant direction in the neighborhood of p. The strength of the local orientation is represented by the magnitude of $\mathbf{d}_s(p)$ and its direction is half of the angle formed by $\mathbf{d}_s(p)$. Since the local orientation angle can varies from 0 to π, doubling the angle gives the range of 2π and thus is efficient in terms of direction continuity (see [Jah95] for more details). The approach of Haglund et. al. is an approximation of the steerable filters since it can recover only the energy response along the "local orientation". However, its calculation is somewhat simpler. Given a local orientation $\mathbf{d}_s(p)$, the directional energy along θ can be approximated by projecting $\mathbf{d}_s(p)$ onto a unit vector $n_\theta = (cos(2\theta), sin(2\theta))$:

$$E_{s,\theta}(p) = \mathbf{d}_s(p) \cdot n_\theta$$

With this approach the Hilbert transforms of the steerable filters $W_{s,d}$ are not necessarily steerable. Haglund suggested using a set of *polar separable* filters whose transfer functions

can be separated into angular parts and radial parts. In the frequency domain (u_1, u_2) this type of filters are conveniently expressed in polar coordinates:

$$\hat{W}_{s,d} = \hat{S}_s(q)\hat{D}_d(\theta)$$

where the notation $\hat{X}$ denotes the Fourier transform of X. The radial part $\hat{S}_s(q)$ is a function of $q = \sqrt{u_1^2 + u_2^2}$ and is in practice a band pass filter that can be implemented by the Laplacian Pyramid scheme. The angular part $\hat{D}_d(\theta)$ is a function of $\theta = \tan^{-1}(u_2/u_1)$. Haglund suggested the use of a set of quadrature pairs with the following angular parts:

$$\begin{aligned} \hat{W}_{s,d}^e &= \hat{S}_s(q)\cos^{2l}(\theta - \theta_d) \\ \hat{W}_{s,d}^o &= \hat{S}_s(q)i\cos^{2l}(\theta - \theta_d)signum(\cos(\theta - \theta_d)) \end{aligned}$$

where $\theta_d = \frac{\pi d}{m}$ for $d = 0, \cdots, m-1$. The indices e and o denote the even and odd components of the quadrature pair. The parameter l determines the angular resolution of these filters. It is necessary that $m > l + 1$ (see [Jah95]). Note that the even component filters are steerable while the odd components are not. If we represent the local orientation $\mathbf{d}_s(p)$ as a complex number whose magnitude corresponds to the strength measure and its angle is double the local direction then the local orientation can be calculated by the following sum (see [Jah95] for a proof):

$$\mathbf{d}_s(p) = \sum_d e^{i2\theta_d} E_{s,d}(p) \quad \text{where} \quad E_{s,d} = (I * K_{s,d}^e)^2 + (I * K_{s,d}^o)^2$$

An example of steerable filters with their Hilbert transform for a single scale is given in Figure 9. The decomposition of the frequency domain using these filters for $m = 3$ is illustrated in Figure 10. The division into rings represents the multi-scale decomposition, while the devision into sectors represents the multi-orientation decomposition.

We use the steerable filters in our demosaicing process in order to calculate local orientations. The results of applying three sets of quadrature steerable filters (three different scales) to a naive interpolation (bilinear) of the raw chair image is presented in Figure 3. From these outputs it is possible to calculate the local orientations as explained. The calculated local orientations for three scales are shown in Figure 11.
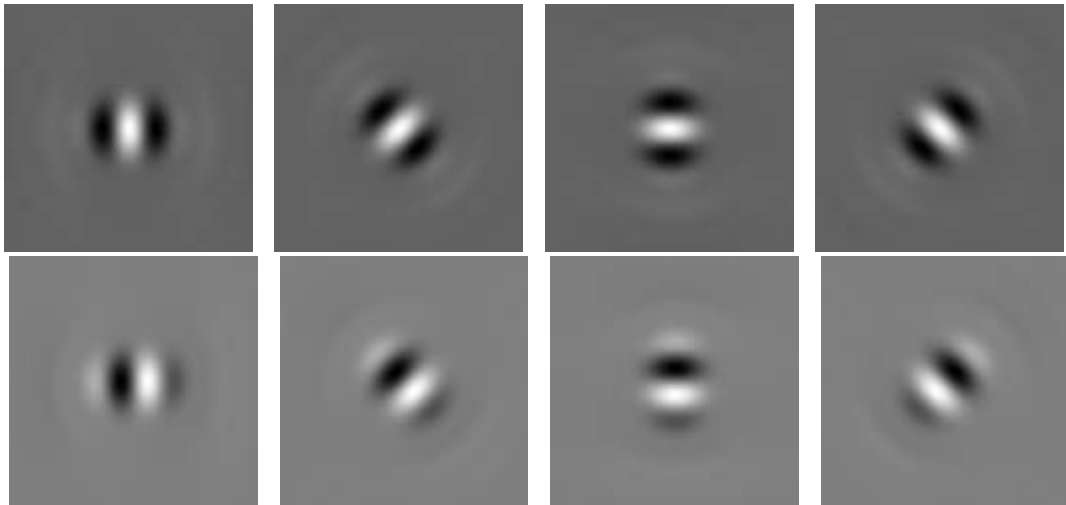
18

Figure 9: Four kernels of steerable filters (top row) and their Hilbert transforms (bottom row) of the 4th octave.
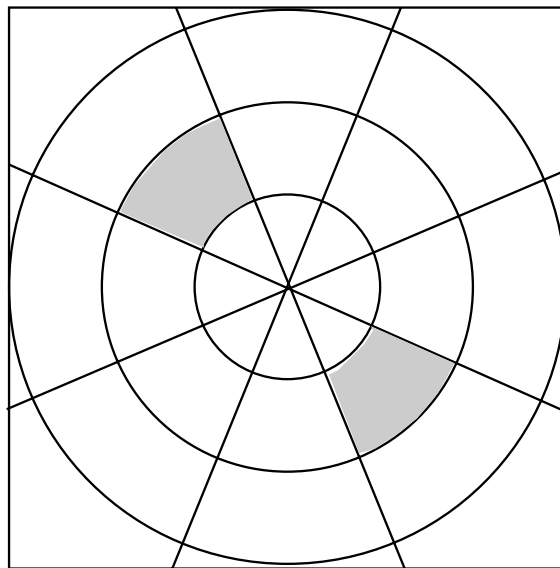

Figure 10: Multi-scale and multi-orientation decomposition of the frequency domain performed using the Steerable Wavelets. The division into rings represents the multi-scale decomposition and the division into sectors represents the multi-orientation decomposition. The grey areas represent a particular filter with a preferred direction of 135°.
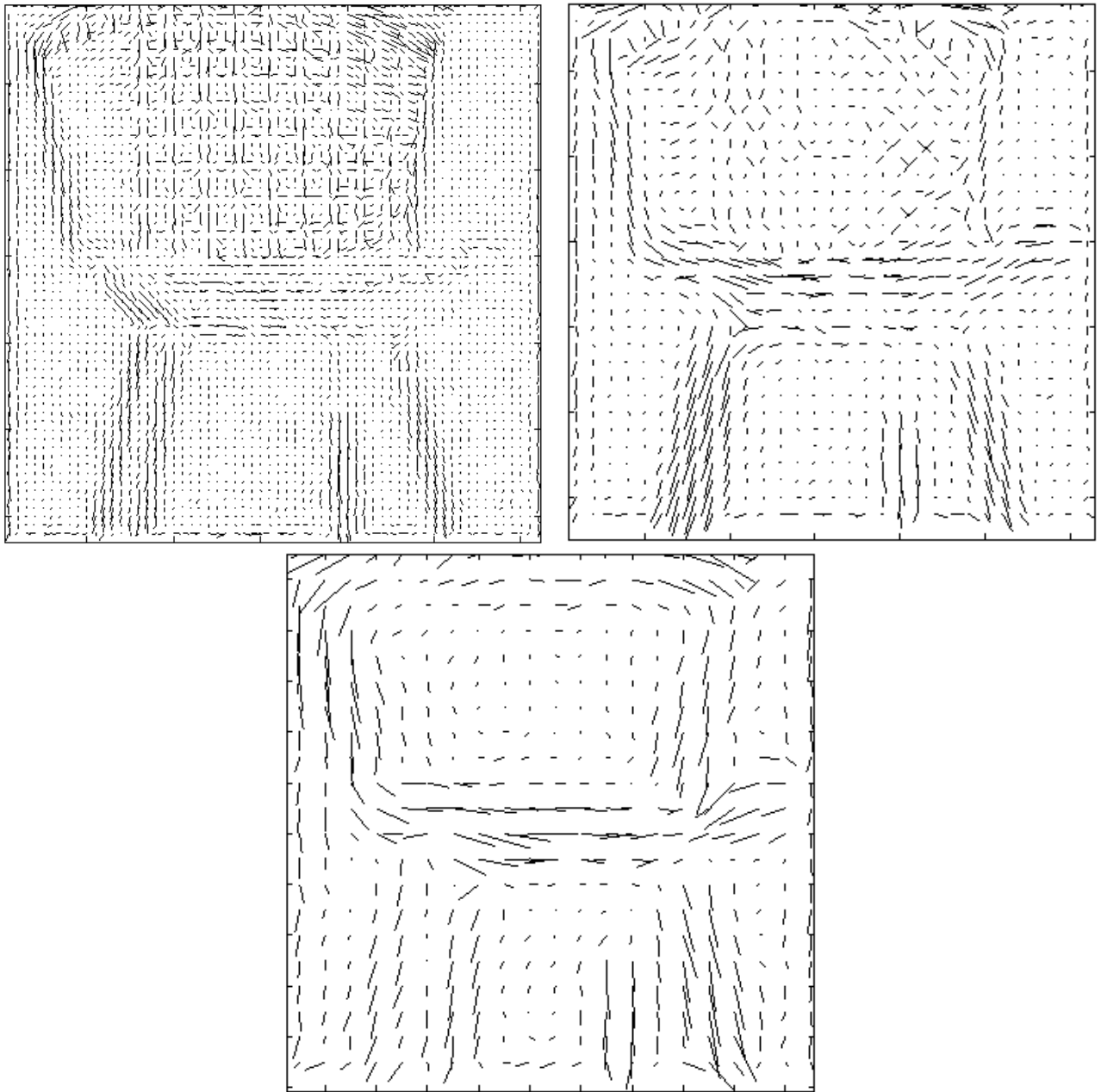
Figure 11: Flow diagrams of the local orientation calculated for three scales of the chair image. At each pixel the direction and the strength of the local orientation is represented by an oriented line segment.

4 Demosaicing of Color Images

At a first glance it would seem that demosaicing of color images can be performed by applying a directional smoothing to each color band (red green and blue) separately. The directional energy should be calculated once, for example, from the luminance image, and directional smoothing of each band can be implemented accordingly. With this approach, the correlation between the color bands is expressed by the fact that a single directional energy map is used. Unfortunately, this approach does not give satisfactory results: Smoothing along dominant orientations indeed retains the sharpness of edges, however, the location of these edges might be subject to small shifts. These small shifts are due to the mis-registration of the color samplings. Misalignment of edges along the color bands results in color artifacts in the demosaic image. This behavior is demonstrated in Figure 12. Figure 13 shows the demosaicing result of the chair image after directional smoothing of each color band separately. Color artifacts are noticeable.

To overcome this problem, a possible solution is to enforce a smooth structure in the color values. However, care should be taken so that edges and lines in the image will not be affected by this smoothing. An efficient color representation space in which color smoothing can be applied is the luminance-chrominance color representation. This representation can be obtained by linearly transforming the RGB values into YIQ values using the following transformation:

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.253 & 0.312 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

In the YIQ representation the Y component represents the luminance (intensity) of the image and the IQ components represent the chrominance. Such a representation is appropriate for our task due to two reasons:

- The luminance-chrominance is an uncorrelated color representation and as such enables treatment of each color band separately [Wan95].

- The human visual system is more sensitive to high frequencies in the luminance band (Y) than in the chrominance bands (IQ). In fact, in the chrominance bands, the human
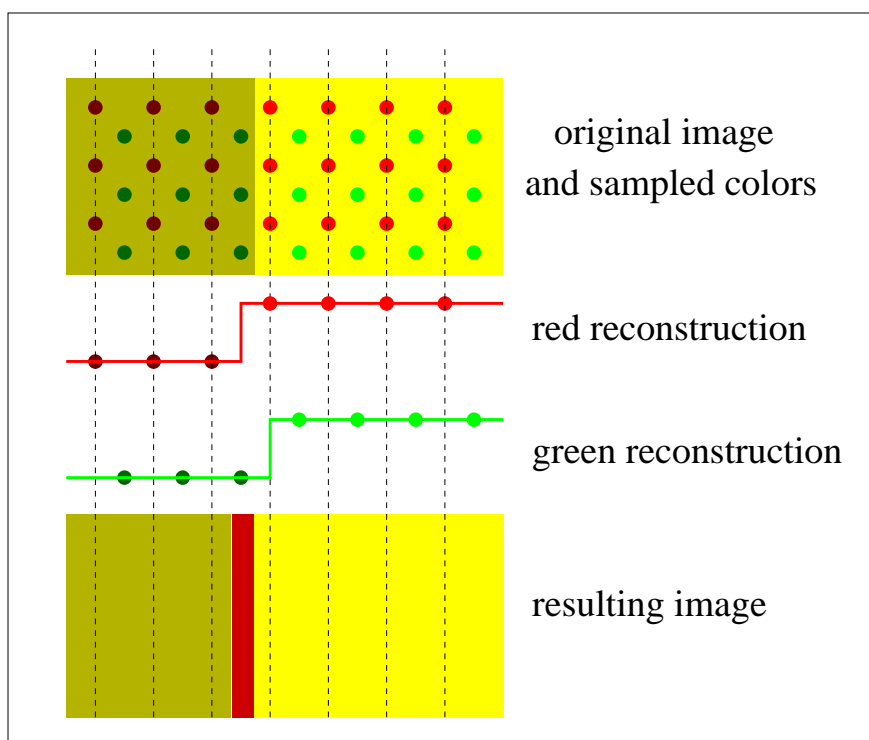
Figure 12: Top: A two-band color image with the corresponding color samples. Middle: The cross sections of the interpolated color bands. The reconstructed steps are sharp due to the directional smoothing. Bottom: A stripe of color artifact due to misalignment of the step edges in the interpolated bands.

visual system is only sensitive to low frequencies. This is shown in Figure 14 (from [PW93]).

Since edges, lines, and textures are composed of high frequencies, their appearance is influenced mainly by the luminance component. This phenomena enables us to enforce a smooth structure in the chrominance bands while preserving high frequency structures in the luminance band. The updated algorithm for demosaicing will include two pathways. A given image is now transformed into a luminance band and two chrominance bands. The luminance band is directionally smoothed according to the algorithm described in Section 2.2. The chrominance bands, however, are isotropically smoothed by a convolution with a Gaussian kernel. The resulting values are then transformed back into the RGB representation and the original samples are reset to their original values. This process is performed several times iteratively until a satisfactory result is produced. In our simulations three to

Figure 13: A demosaic image after applying the directional smoothing on each color band separately. Note the color artifacts near edges.

four iterations produce good results. The suggested algorithm is illustrated in Figure 15.

5   Results

The proposed algorithm was tested on a wide variety of images. Some of them were acquired by the Kodak digital camera (DCS-200) where raw data was provided as well as the demosaic images. These images supply an objective comparison between existing demosaicing algorithm (Kodak's) and the suggested method. The rest of the test images were generated from given RGB images where raw data was produced by taking a partial sampling of the original values, simulating a digital camera. Figures 16, 17 and 18 are three examples of demosaic images produced by the suggested algorithm, along with the results provided by the Kodak camera. It can be seen that the Kodak algorithm produces color artifacts mainly along edges and in textured areas (the chair and grass areas in Figures 16, 17, and the wood pattern in Figure 18). These artifacts were drastically reduced using the suggested approach.

Figures 19 and 20 are two examples of demosaic images along with the original RGB images.
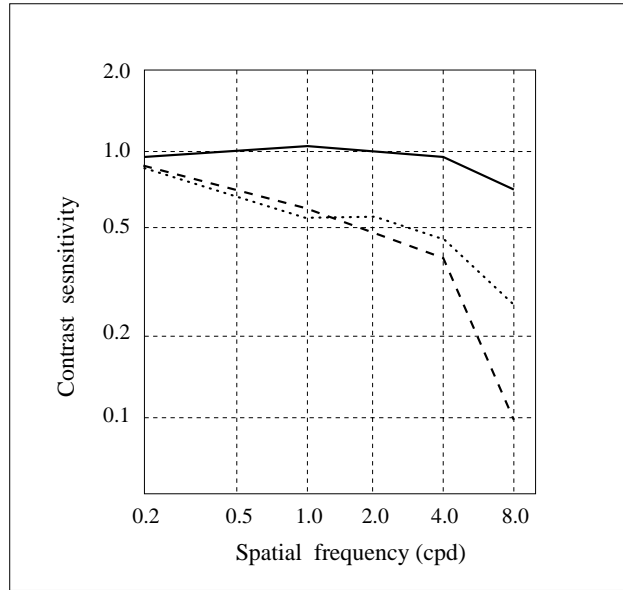
Figure 14: Luminance-chrominance sensitivity of the human visual system with respect to spatial frequency. The solid line represents the luminance sensitivity while the dashed and the dotted lines represent the chrominance sensitivities. Taken from [Poirson and Wandell 1993].

Indeed, color artifacts were reduced, however, due to the lower sampling rate, the demosaic images are slightly blurred compared to the originals. Image 21 is a good example demonstrating the limitation of the suggested approach and actually of any demosaic algorithm. Due to the perspective warp in this image, the fence rungs of the lighthouse form a monotonically increasing frequency pattern. At a particular frequency, the demosaic algorithm breaks down and produces a strong color aliasing. Notice that the aliasing artifacts occur in the area where the rungs falls on every other pixel which is exactly the Nyquist frequency of this image. This demonstrates that the demosaic algorithm cannot recover high frequencies above the Nyquist limit.

6 Computational Aspects

The computation cost of the proposed algorithm is composed of three parts:

1. The transformation of the RGB image into YIQ representation.

2. The decomposition of the luminance band into the steerable wavelets representation.
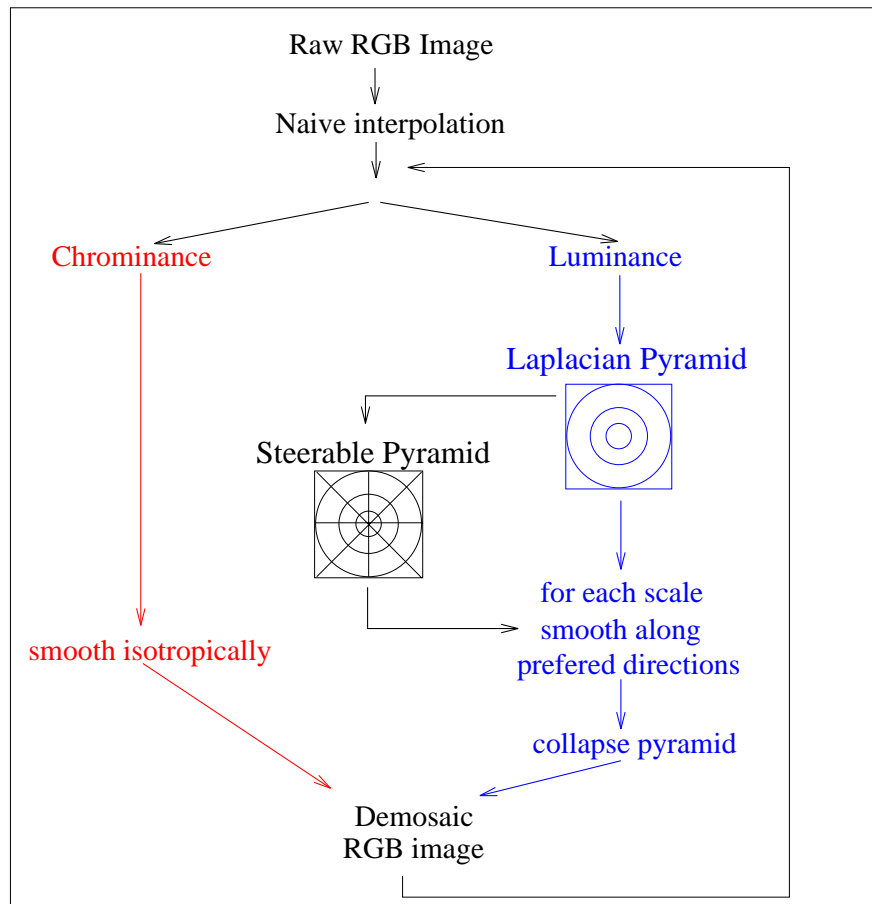
Figure 15: Demosaicing algorithm of color images.

This includes the application of two filters per each directions (the quadrature pair) and a repetition of the directional decomposition for each scale level.

3. The directional smoothing of the luminance component and the isotropic smoothing of two chrominance components.

This complexity can be drastically reduced if some approximations are performed. First, since the demosaic color artifacts take place mainly in the high frequencies, it is possible to work in the highest scale level or even in the original image space. This, of course, will not perform well in the low frequencies domain, however, we assume that low frequencies are interpolated quite accurately. Next, the number of directional decomposition can be reduced. We used a four directional decomposition, which required the application of eight filters,
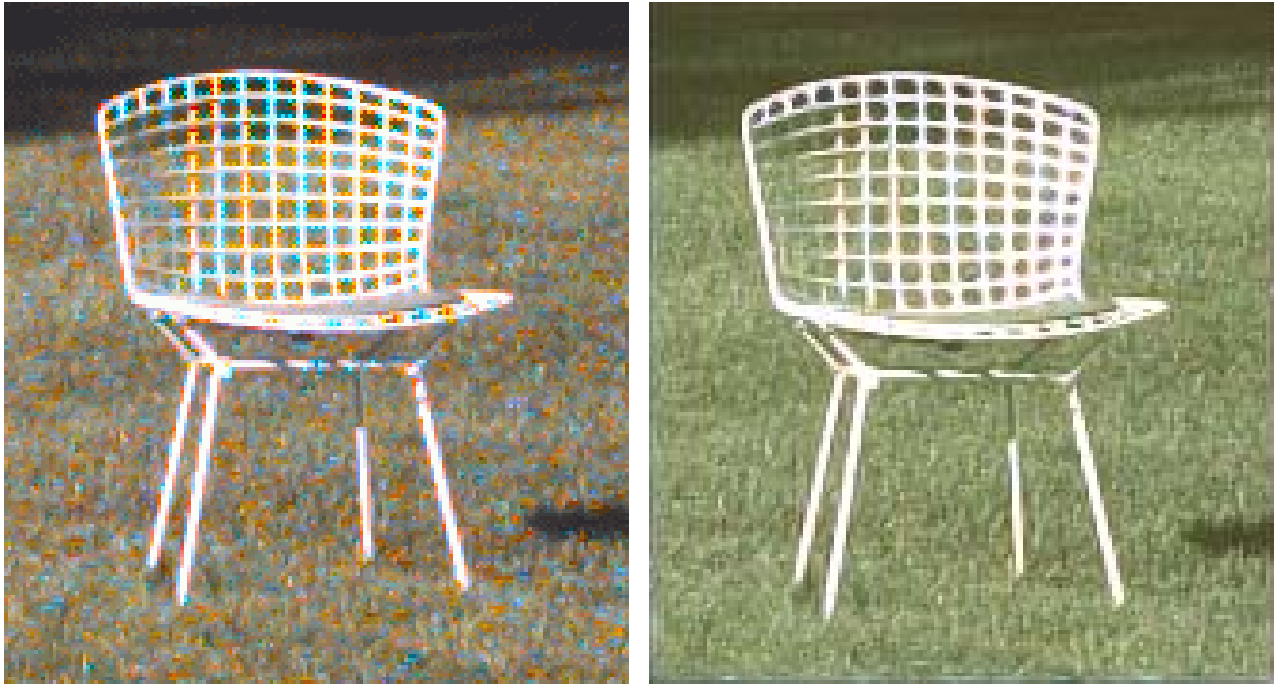
Figure 16: Demosaic results of color images. Left: The demosaic result of the chair image as produced by the Kodak camera algorithm. Note the color artifacts near the chair rungs and in the grass area. Right: The demosaic result using the suggested approach.

however, three directional decomposition gives reasonable results as well. Finally, since the RGB to YIQ transformation and the smoothing operation are both linear operations, it is possible to propagate the smoothing operation into the RGB space. This will eliminate the need to transform from RGB to YIQ and visa versa at each iteration. Additionally, working in the RGB space can avoid the calculations of pixel values existing in the original sampled image.

7   Conclusion

A new demosaic algorithm is proposed which interpolates a color image given its partial sampling from a digital camera. The technique is based on the steerable wavelets decomposition where image values are interpolated in their appropriate scale and in their appropriate direction. The interpolation is performed in the luminance-chrominance color representation. Chrominance values are interpolated using the classical regularization approach. Luminance values are interpolated using a directional regularization scheme. This approach gives a

faithful interpolation of edges and textured areas while reducing color artifacts.

Acknowledgments

Figure 17: Demosaic results of color images. Top: The demosaic result as produced by the Kodak camera algorithm. Bottom: The demosaic result using the suggested approach.
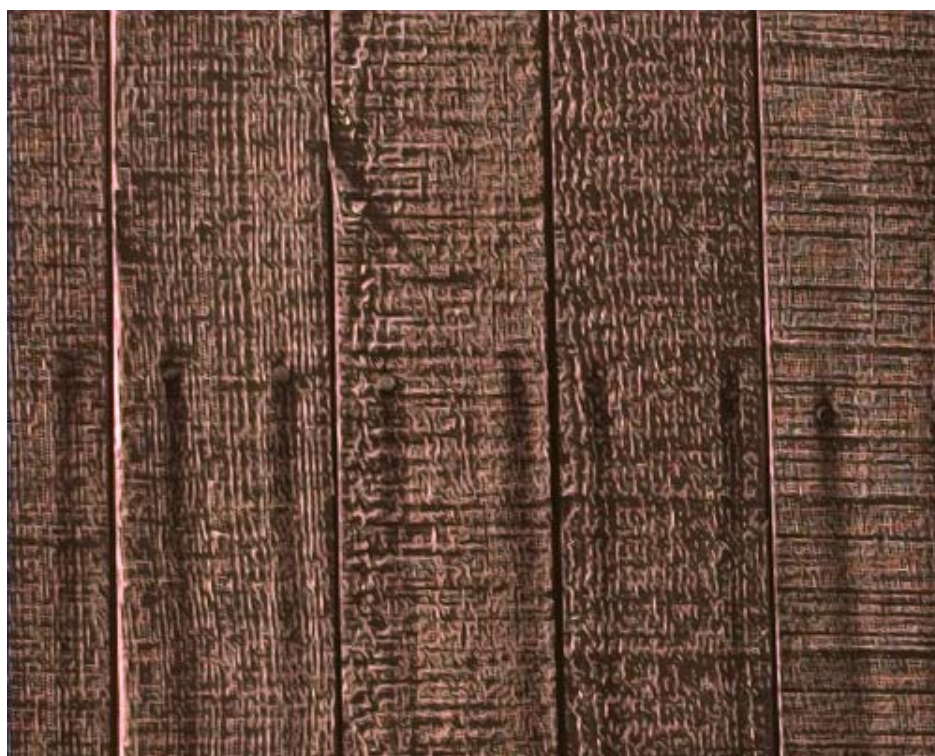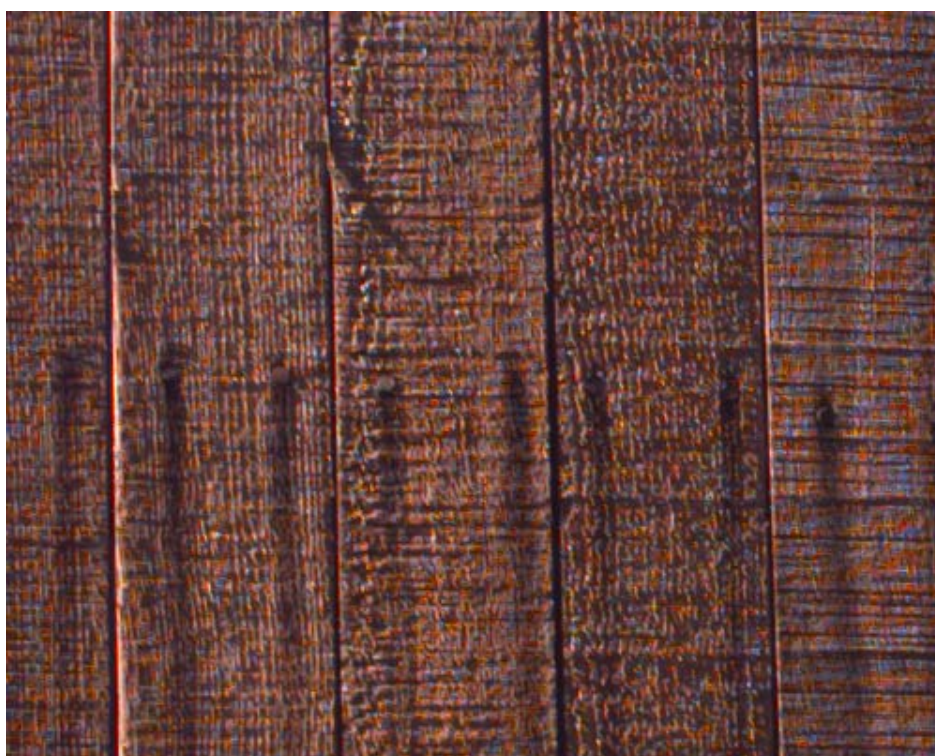
Figure 18: Top: The demosaic result as produced by the Kodak camera algorithm. Bottom: The demosaic result using the suggested approach.

Figure 19: Left: The original RGB image. Right: A demosaic result using a partial sampling of the original image.

Figure 20: Top: The original RGB image. Bottom: A demosaic result using a partial sampling of the original image.

Figure 21: Left: The original RGB image. Right: A demosaic result using a partial sampling of the original image.

8 References

[BA83]   P.J. Burt and E.H. Adelson. The laplacian pyramid as a compact image code. *IEEE Trans. Communications*, 31(4):532–540, 1983.

[BA91]   J.R. Bergen and E.H. Adelson. Theories of visual texture perception. In D. Regan, editor, *Spatial Vision*. CRS press, 1991.

[BR96]   M. Black and A. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision*, 19(1):57–91, 1996.

[Bra78]   R.N. Bracewell. *The Fourier Transform and its Applications*. McGraw-Hill, 1978.

[Bra94]   D.H. Brainard. Bayesian method for reconstructing color images from trichromatic samples. In *Proc. of the IS&T /ICPS*, 1994.

[BS95]   D.H. Brainard and D. Sherman. Reconstructing images from trichromatic samples: From basic research to practical applications. In *Proc. of the IS&T /SID*, 1995.

[BSmH97]   M. Black, G. Sapiro, D. marimont, and D. Heeger. Robust anisotropic diffusion. *Submitted*, 1997.

[BZ87]   A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, 1987.

[FA90]   W.T. Freeman and E.H. Adelson. Steerable filters for early vision, image analysis, and wavelet decomposition. In *International Conf. on Computer Vision*, pages 406–415, 1990.

[FA91]   W.T. Freeman and E.H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(9):891–906, 1991.

[Fie93]   D. Field. Scale-invariance and self-similar wavelet transforms: an analysis of natural scene and mammalian visual systems. In *Wavelets, Fractals, and Fourier Transforms*. Eds. Oxford University Press, 1993.

[GG84]   S. Geman and D. Geman. Stochastic relaxation, gibbs distribution, and the bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[HK84]   B.R. Hunt and O. Kubler. Karhunen-loeve multi-spectral image restoration. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 32(3):592–599, 1984.

[HKG93] L. Haglund, H. Knutsson, and G.H. Granlund. Scale and orientation adaptive filtering. In *Proceedings of the 8th Scandinavian Conference on Image Analysis*, pages 847–855, 1993.

[Jah95] B. Jahne. *Digital Image Processing, Concepts, Algorithms, and Scientific Applications*. Springer-Verlag, 1995.

[Ker87] D. Kersten. Predictability and redundancy of natural images. *Journal of the Optical Society of America*, 4:2395–2400, 1987.

[Ker96] D. Keren. An adaptive bayesian approach to demosaicing color images. Technical Report HPL-96-129, Hewlett-Packard, 1996.

[MMP87] J. Marroquin, J. Mitter, and T. Poggio. Probabilistic solution of ill-posed problems in computational vision. *Journal of the Optical Society of America*, 82(397):76–89, 1987.

[MMRK91] P. Meer, D. Mintz, A. Rosenfeld, and D.Y. Kim. Robust regression methods for computer vision: A review. *International Journal of Computer Vision*, 6(1):59–70, 1991.

[OF96]   B.A. Olshausen and D.J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural image. *Nature*, 381:607–609, 1996.

[Pm90]   P. Perona and J. malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[PW93] A.B. Poirson and B.A. Wandell. The appearance of colored patterns: Pattern-color seperability. *Journal of the Optical Society of America*, 12:2458–2471, 1993.

[SFAH92] E.P. Simoncelli, W.T. Freeman, E.H. Adelson, and D.J. Heeger. Shiftable multi-scale transforms. *IEEE Trans. Information Theory*, 38(2):587–607, 1992.

[SGVE89] M.S. Silverman, D.H. Grosof, R.L De Valois, and S.D. Elfar. Spatial frequency organization in primate striate cortex. *Proc. Natl. Aca. Sci. U.S.A.*, 86(2):711–715, 1989.

[Wan95] B. Wandell. *Foundations of Vision*. Sinaur Associates, 1995.