Demosaicing using Artificial Neural Networks

Oren Kapah[a] and Hagit Z. Hel-Or[b]

[a]Dept of Computer Science, Bar-Ilan University, Ramat-Gan 52900, Israel
[b]Dept. of Computer Science, Haifa University, Haifa 31905, Israel

ABSTRACT

The problem of color image enhancement and the specific case of color *demosaicing* which involves reconstruction of color images from sampled images, is an under-constrained problem. Using single-channel restoration techniques on each color-channel separately results in poorly reconstructed images. It has been shown that better results can be obtained by considering the cross-channel correlation. In this paper, a novel approach to demosaicing is presented, using learning schemes based on Artificial Neural Networks. Thus the reconstruction parameters are determined specifically for predefined classes of images. This approach improves results for images of the learned class, since the variability of inputs is constrained (within the image class) and the parameters are robust due to the learning process. Three reconstruction methods are presented in this work. Additionally, a selection method is introduced, which combines several reconstruction methods and applies the best method for each input. The first reconstruction method presented in this work is the Perceptron method which considers linear restoration models. Reconstruction using the Perceptron method yields excellent results in low-frequency regions of the image. However, in the high-frequency regions (especially in desaturated color regions) the results of the Perceptron network are not satisfactory. This is shown to be due to the linearity and the uniformity of the Perceptron network. Restoration using the Backpropagation method yields relatively good results in high-frequency regions. However, in the low-frequency regions the results are not satisfactory since it fails to reconstruct the correct color of the region. The Selector method is shown to be useful in combining both of the previous methods in one reconstruction scheme which works in both cases of low and high frequency regions. The Quadratic Perceptron method is a new development extending the Perceptron network in order to overcome the limitations of the latter network and provide a better solution for the reconstruction in the high-frequency regions.

Keywords: Demosaicing, Artificial Neural Networks, Image Restoration, Enhancement, Color Images

1. INTRODUCTION

Images are typically represented by two dimensional matrices of pixels. In color images, the color of each pixel is usually represented by a triplet (R,G,B) which specify the red, green and blue values. This representation however, is not available directly from the output of many color CCD cameras. Rather, these cameras provide only a single photo-sensor response at each pixel. The camera contains several classes of photo-sensors, each class has a distinct spectral response. Usually there are three classes, referred to as red, green and blue, corresponding to the three color bands of the digital image. There exist several designs of Color Filter Array (CFA) patterns, but the most successful, and thus, the most commonly used is the Bayer CFA.[1] Figure 1 provides a schematic illustration of this design.

In order to obtain a full color image, the values of the two missing color channels at each point must be reconstructed. This process of reconstruction is known as *demosaicing.* The most straightforward approach to demosaicing is to apply one of the standard reconstruction methods to each color band separately. The disadvantage of this approach is that the phase difference between the sampling of different color bands may cause the appearance of new colors in the image, particularly around the edges, and thus, distorting the image. This kind of distortion is illustrated in Figure 2 for a 1D example. A black-white edge is captured by the CCD array as shown in Figure 2a. Assuming a simple linear interpolation on each band separately (Figure 2b and Figure 2c) and assuming a perfectly reconstructed blue channel, a distortion is introduced in the reconstruction (Figure 2d).

This type of distortion is very strong in reconstructed images when there are many edges in the same area. This problem originates from the fact that the resolution of each band is limited by the Nyquist sampling rate of the specific band which is lower than the sampling rate of the full image sampling rate. Thus when reconstructing each

Send correspondence to H.Z. Hel-Or. E-mail: hagit@cs.haifa.ac.il

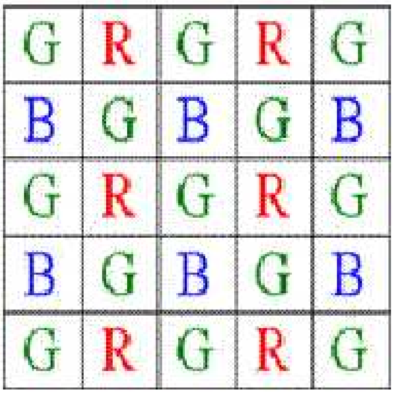

Figure 1. Bayer Color Filter Array

band separately, since the information from the other bands and the correlation between them are not taken into consideration, a poor reconstruction is obtained. A better reconstruction of the image can be obtained if we take into consideration the cross-channel correlation between the different color bands. Recent work demonstrated the importance of using the cross-channel correlation information in demosaicing methods. Unfortunately, the algorithms suggested still produce some artifacts in the reconstructed images. In this work, a new approach is suggested for demosaicing, based on learning from examples.

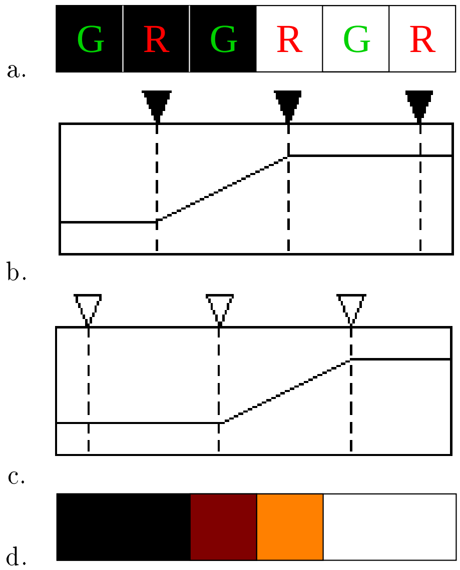

Figure 2. (a) The original image as sensed by the CCD. (b) The reconstructed red channel. (c) The reconstructed green channel. (d) The reconstructed image.

2. PREVIOUS APPROACHES

The demosaicing problem is a special case of the image reconstruction and enhancement problem. The most straightforward approach for solving the demosaicing problem in color images, is to apply one of the standard reconstruction methods for gray-scale images on each channel separately. Many methods have been proposed for single channel reconstruction: Interpolation, Regularization and Inverse Filtering to name a few (see[2,3] for a review).

As discussed above, reconstructing each channel separately produces artifacts. Better reconstruction of the image can be obtained by taking into consideration the cross-channel correlation. Several methods have recently been suggested to solve the demosaicing problem using cross-channel correlation.

In[4] the Green channel is first reconstructed based on classification of local neighborhoods into pattern classes, then the red and blue channel are reconstructed based on the reconstructed Green channel. An improvement was suggested in[5] where a weight function is used in the reconstruction stage to predict the local structure of the image. Additionally, an enhancement stage was added following the reconstruction which is based on reaction-diffusion techniques. Another method was introduced in[6,7]in which the reconstructed image is chosen to best match: (a) the image obtained by the camera CCD and (b) some prior knowledge about images in the real world. In order to achieve this, Bayesian theory is applied. A different approach was developed at the Polaroid Corporation.[8] Their algorithm takes into account the fact that the optical system and the spectral characteristics of the color filters, introduce spatial and color correlations into the sampled image. The algorithm measures the characteristics of the Color Filter Array (CFA), and computes the best coefficients for the linear combination of the neighboring pixels which produce the reconstruction. Then, for each acquired image, the full image is reconstructed using the calculated coefficients. In[9] an algorithm is proposed for removing the artifacts in the reconstruction. The algorithm is based on the fact that unsynchronized edges in two channels produce a spike in the difference image of these two channels. In order to remove the spikes in the color difference image without blurring the image, the median filter is used. Another approach[9] is to use the above technique in an angle based approach. The main idea of this algorithm is to consider the color coordinates at each pixel, as three dimensional vectors in a spherical coordinate system. Unsynchronized edges in two channels then produce a spike in the angle difference image.

Previous approaches that take into consideration the inter-channel correlations, improve performance over standard channel-independent reconstruction. However, there is room for improvement. Template matching, filtering and luminance based approaches, which are the basis for the various techniques mentioned above, are limited in the capability of reconstructing a wide variety of color images with different spatial and chromatic characteristics.

In this paper a new approach is proposed and tested. The demosaicing parameters are determined specifically for predefined classes of images. These parameters are obtained through a learning process using artificial neural networks. This approach improves results for images of the learned class, since the variability of inputs is constrained (within the image class) and the parameters are robust due to the learning process. Details are given in the following sections.

3. THE NEURAL NETWORK APPROACH

Artificial Neural Networks(ANN's) provide a robust tool for approximating a target function given a set of input-output examples. Algorithms such as the Backpropagation and the Perceptron use gradient-descent techniques to tune the network parameters to best-fit a training set of input-output examples. For certain types of problems, such as, learning to interpret noisy outputs of sensors, artificial neural networks are among the most effective learning methods. The ANN learning methods consist of two stages: (1) the Learning stage (2) the Fixed stage. In the learning stage, the weights of the edges in the net are tuned to best-fit the learning set of examples in which the inputs and the target outputs are given. In the fixed stage, the weights are fixed and used to calculate the estimated output for any input. For more details and a background on ANN see.[10]

The motivation in using ANN for solving the demosaicing problem, is based on the fact that this problem does not have one absolute mathematical solution. The missing values can not be reconstructed using only mathematical analysis because the problem is under-constrained. There are many images which are in accord with the sampling rule as shown in the previous chapters. Thus, some assumptions must be incorporated in order to fill in the missing values. These assumptions might not necessarily be correct for all cases and may lead to errors in the reconstruction of new images. Additionally, there might be other constraints that exist in the real world that are not taken into account in the reconstruction process and taking advantage of these constraints may result in better reconstruction of the images.

The ANN technique provides a robust tool for learning the reconstruction function from a class of images. The information regarding the correlation between the different color bands are extracted from the example images, and are expressed by the learned weights. Using the ANN technique, this information can be exploited without the need of formulating it. In practice, better reconstruction can be achieved by training several ANN's, each to a different

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

Figure 3. The inputs to the implemented networks

class of images in order to simplify the reconstruction. In this way, specific attributes of each class of images can be exploited to improve the reconstruction.

To apply the ANN methods on the demosaicing problem, a set of learning examples must be provided. Each example contains an input which is a sampled image obtained by a camera CCD and an output which is the desired reconstructed full color image. Since the output images of the CCD cameras are already sampled, and the full RGB image is unavailable, we simulate the CCD sampling process by taking full color RGB images and sampling in accord with the CCD sensor array.

4. USING ANN TO SOLVE THE DEMOSAICING PROBLEM

Several ANN models were used to reconstruct color images of various classes. As will be seen, the different models produce results of varying quality. Common to all models, is the structure of the input-output example pairs. The Bayer CFA is assumed (Figure 1) which consists of a 2×2 pattern that is repeated throughout the CFA. Given that a priori there is no difference in the likelihood of image values appearing in any such 2×2 region in the image, the ANN networks were built to reconstruct any full $2 \times 2 \times 3$ part of the image by reconstructing the missing channel values in this 2×2 matrix. The full reconstructed image is obtained by applying the learned network over all 2×2 patterns in the image. In order to obtain the best reconstruction of the missing values, the local information obtained by the camera's CCD is also supplied as input to the network. Thus, the 2×2 region and their immediate neighbors are supplied as inputs to the implemented networks as illustrated in Figure 3. As a result, the implemented networks consist of 16 inputs supplied as a 4×4 matrix of the sampled values (the 2×2 pattern and the 12 neighboring sampled values), and 8 outputs (2 missing values of each pixel of the 2×2 matrix).

For the experiments, two classes of images were created. The first class contains structured images (buildings with fences, windows, etc) which have high frequency regions containing desaturated colors and the second class contains flower images which have high frequency regions containing saturated colors. Each class of images was divided into two subsets. One set of images was used in the learning procedure and the second set of images was used for testing the implemented network. For the training example set, each image provided several hundred 2×2 input-output pairs.

For each ANN method, the network was trained twice, once for each class of images. The training stage was applied so as to complete 50 loops over the example set. For each network, the weights that minimize the Mean Squared Error (MSE) on the learning set were selected as the learned weights. Given the learned weights, the output of these networks on the test subsets of images was calculated.

4.1. LINEAR RESTORATION - THE PERCEPTRON MODEL

Several algorithms suggested for solving the demosaicing problem are within the restriction of linear restoration. However, either they do not use the spatial and spectral correlation between the channels that exist in real images as in,[8] or they require formulation of some prior knowledge of the features existing in real world images as in[6] which complicates the restoration. The Perceptron model provides us with a method of obtaining simple restoration without the need of specifying any previous knowledge of the spatial and spectral correlation between channels or the probability of images to appear in the real world. This information is learned by the Perceptron from a class of example images and is taken into account in the reconstruction of the image through the values of the coefficients.

a.  b. 

Figure 4. (a) Reconstruction of the Flowers image using channel independent interpolation. (b) The reconstructed Flowers image using the Perceptron network trained over the structures images set.

Another benefit of the Perceptron method is the simplicity of the learned function which makes its performance easier to understand by studying the weights.

We used a Perceptron Network with full connectivity. For details on the structure and training rules of the Perceptron model see.[10,3] In the demosaicing problem a function with a continuous range should be obtained, thus, the threshold stage in the Perceptron units was eliminated. All inputs and outputs were normalized to the range [0..1] (0 corresponds to the lowest sensor value [dark] and 1 corresponds to the highest sensor value [bright]). In order to restrict the outputs to be within the range, these maximum and minimum values were enforced by truncation.

The results of reconstructing a test image from the flowers class is shown in Figure 4 (for additional result examples see[3]).[*]

As can be seen, the results of the Perceptron network are very good in the low frequency regions. In the high frequency regions which contain saturated colors (such as the green leaves) the results are not bad, since, the demosaicing artifacts are not visible there. However, in the high frequency regions which contain desaturated colors (such as the dry leaves), the result of the Perceptron network is not satisfactory, since, the aliasing described in Section 1 appears. This is due to the limitations of the Perceptron network, in which the same weights are used for all the input-colors in the image regardless of whether they arise from a high frequency region or a low frequency region. Since the outputs in the Perceptron network are linear combination of the inputs, the Perceptron network cannot react well to changes in edge positions and directions. This results in slight blurring at the edges which become noticeable in the desaturated color regions.

4.2. NONLINEAR RESTORATION - THE BACKPROPAGATION MODEL

The Backpropagation network, capable of learning complex non-linear functions is expected to produce better reconstruction than the Perceptron network, especially in the high frequency regions of the image. We used a two-layer Backpropagation network with full connectivity. The hidden layer was constructed as a 4×4 matrix connecting the input and output layers. For details on the structure and training rules of the Backpropagation model, see.[10,3]

The result of reconstructing a test image from the structure class is shown in Figure 5.

As can be seen in the reconstructed images, the Backpropagation network performs well in high frequency regions. In fact, due to the sharp slope of the sigmoid around the middle of the range, the values close to the middle of the range are scaled. Thus, small color contrasts at edges in the hidden layer output image are enhanced by the sigmoid function, resulting in higher color contrast in the final reconstructed image. This effect sharpens the reconstructed image. In the low frequency regions, however, the Backpropagation network failed to reconstruct the correct color of the pixels. This problem is due to the nonlinearity of the sigmoid function involved in the Backpropagation process. This nonlinearity contracts extreme pixel values of the rage [0..1] toward the middle of the range. Thus, saturated colors (expressed as extreme values of one or more of the color channels) become desaturated.

---

[*]Throughout this paper only one example is shown due to space limitation. Additional examples in color can be found in[3] and at url: www.cs.haifa.ac.il/hagit/demos/demosaicing and url: www.cs.haifa.ac.il/hagit/papers/spie00.ps

Figure 5. The reconstructed flowers image using the Backpropagation network trained over the flowers images set

4.3. THE SELECTOR MODEL

As described previously, the two implemented ANN models performed differently. The Perceptron method performed well in the low frequency regions of the image and failed in the high frequency regions, while the Backpropagation method performed relatively well in the high frequency regions and failed in the low frequency regions. This observation implies that better results can be obtained by applying the appropriate method locally according to local characteristics of the image. Rather than selecting the appropriate reconstruction method interactively, a new artificial neural network was built to determine, for each set of inputs, which method to apply in order to obtain the best reconstruction.

The *Selector* method is thus an ANN which is supplied with a set of reconstruction methods (in our case the Perceptron network and the Backpropagation network) and is trained to associate between inputs and a discrete output representing the index of the reconstruction method that should be used. An illustration of the Selector model is shown in Figure 6.

The motivation for using the Selector method is that although the reconstruction function is difficult to find or compute, the selection function may be very simple because the selection function does not have to provide exact values of outputs, instead it need only provide a discrete output within a small range of values.

The Selector network was chosen to be a Backpropagation network so as not to restrict the selection function to linearly separable functions. However a slight modification was incorporated into the classic Backpropagation structure in order to obtain better results. This modification enables the learning procedure to take into account an

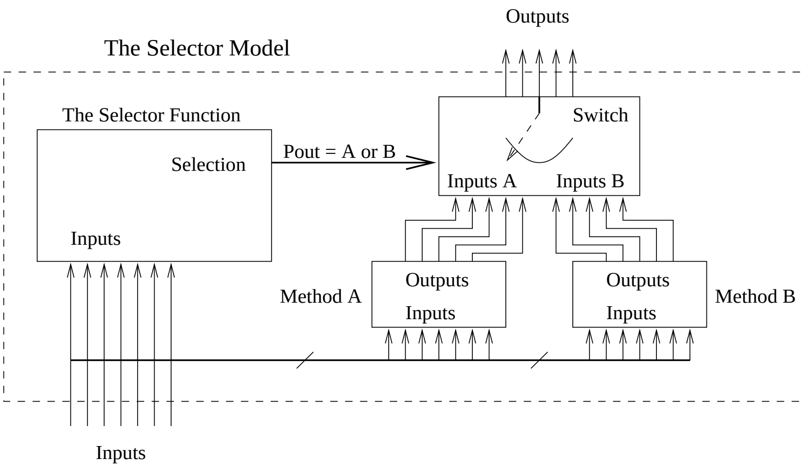

Figure 6. The Selector model

a. b.

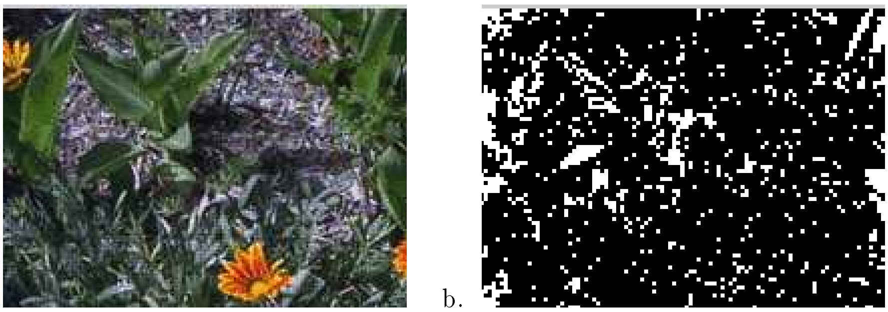

Figure 7. (a) The reconstructed Flowers image using the Selector network trained over the flowers images set. (b) The selection image resulting from applying the Selector on the Flowers image. Black represents locations for which the Backpropagation method was selected and white represents locations for which the Perceptron method was selected.

importance factor associated with each training example. There are input examples for which there is only a small difference between the reconstruction quality obtained using the (two) available reconstruction methods. Thus, it is not critical, to choose the best method for these cases. While in other examples there might be a big difference between the reconstruction results and choosing the best method is important. in practice, the importance factor determines the influence on the network weights during the training stage. Additional details on the Selector method can be found in.[3]

The Selector method can be easily extended to choose from among any finite set of restoration functions. One must consider, however, that the greater the number of restoration functions given to the Selector, the more complex becomes the selection function and the efficiency of the Selector method decreases. Thus, when using this method with several restoration functions it is important to reduce the number of functions to a minimum.

In our experiments, both the Perceptron and the Backpropagation reconstruction methods were used with the weights that were learned over the two training sets mentioned above. As in the previous methods, the training stage involved pairs of input-output examples. In the Selector method, for each input example, the two reconstruction methods were applied and the results were compared with the target output. The target of the selection function was set to be the index of the method providing reconstruction results having the smallest MSE with respect to the target output. The importance factor of each example was set to equal the absolute difference between the two MSE values.

The result of the Selector method on a test image of the Flowers class is shown in Figure 7a. The reconstruction method chosen for each pixel in the image is shown in Figure 7b. In the these images, black represents locations for which the Backpropagation method was selected and white represents locations for which the Perceptron method was selected. It is not surprising that the Perceptron method was selected in the low frequency regions of the image and the Backpropagation method was selected in the high frequency regions.

4.4. QUADRATIC RESTORATION - THE QUADRATIC PERCEPTRON MODEL

As shown above, the Perceptron network performs well in low frequency image regions. However, it is limited in high frequency regions, due to the fact that the same weights are used for all the input-vectors in the image regardless of whether they arise from high frequency regions or low frequency regions.

The *Quadratic Perceptron* was developed to meet the requirements of different weights for different inputs. In the Quadratic Perceptron the weights are computed for each input vector as a function of the inputs. In the training stage, instead of learning constant weights, the functions which produce the best weights for each input vector are learned. The Quadratic Perceptron is implemented as a single-layer Perceptron network. However, associated with each weight of the network, is an additional Perceptron subnetwork which is trained to evaluate the weights function. In order to ensure maximum accuracy, all the inputs to the Quadratic Perceptron network were used as inputs to each of the Perceptron subnetworks used to compute the weights w_i. That is, the weights are a linear function of the inputs and are computed for each input vector, using a different Perceptron subnetwork for each weight. Thus in

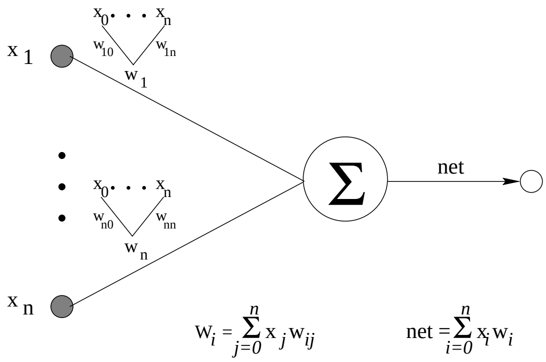

Figure 8. The Quadratic Perceptron unit model

practice the output is a quadratic function of the inputs. In order to enable a constant term in the computation of the weights, a constant input x_0 with value 1 and weight w_0 is added to each Perceptron subnetwork. The Quadratic Perceptron network model is shown in Figure 8. For more details see.[3]

The result of reconstructing a test image from the structure class is shown in Figure 9.

As can be seen the Quadratic Perceptron provides good results in the high-frequency regions (as in the dry leaves) as well as in the low-frequency regions (as in the green leaves). The overall performance of the Quadratic Perceptron network is the best from amongst the methods presented in this work.

Figure 9. Reconstructed Flowers image using the Quadratic Perceptron network trained over the Flower images set

5. CONCLUSIONS

In this paper we present a novel approach to solving the demosaicing problem, namely, using learning schemes. Several approaches were presented using Artificial Neural Networks, including the Perceptron, the Backpropagation, the Selector and the newly developed Quadratic Perceptron Network. Additionally the Selector method was presented which is capable of combining the advantages of several given reconstruction methods.

The demosaicing parameters are determined specifically for predefined classes of images. This approach improves results for images of the learned class, since the variability of inputs is constrained (within the image class) and the parameters are robust due to the learning process. The ANN techniques provide a robust tool for learning the reconstruction function from a class of images. The information regarding the correlation between the different color bands are extracted from the example images, and are expressed by the learned weights. Using the ANN technique, this information can be exploited without the need of formulating it. In practice, better reconstruction is achieved by training several ANN's, each to a different class of images. In this way, specific attributes of each class of images can be exploited to improve the reconstruction.

REFERENCES

1. B. E. Bayer, "Color imaging array." United States Patent 3,971,065, 1976.
2. H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice Hall, Inc., 1977.
3. O. Kapah, "Image restoration solutions for digital cameras using artificial inteligent networks," Master's thesis, Bar-Ilan University, June 1999.
4. D. R. Cok, "Reconstruction of CCD images using template matching," *IS&T's 47th Annual Conference/ICPS: The Physics and Chemistry of Imaging Systems* **2**, pp. 380–385, 1994.
5. R. Kimmel, "Demosaicing: Image reconstruction from color CCD samples," Tech. Rep. LBNL-40820, UC Berkeley, September 1997.
6. D. H. Brainard, "Bayesian method for reconstruction color images from trichromatic samples," *IS&T's 47th Annual Conference/ICPS: The Physics and Chemistry of Imaging Systems* **2**, pp. 375–380, 1994.
7. D. H. Brainard and D. Sherman, "Reconstruction images from trichromatic samples: From basic research to practical applications," in *Third IS&T/SID Color Imaging Conf.: Color Science, Systems, and Application*, pp. 4–10, 1995.
8. M. A. Wober and F. R. Soini, "Method and apparatus for recovering image data through the use of a color test patterns." United States Patent 5,475,769, 1995.
9. M. Gamer, "Using cross-channel correlation to restore multispectral images," Master's thesis, University of Haifa, July 1997.
10. T. M. Mitchell, *Machine Learning*, The McGraw-Hill Companies, Inc., 1997.