RECONSTRUCTION OF SPATIAL AND CHROMATIC INFORMATION FROM THE CONE MOSAIC

DAVID ALLEYSSON
Laboratoire de Psychologie et NeuroCognition LPNC, CNRS UM5105
Université Pierre-Mendès France,
Grenoble, France

BRICE CHAIX DE LAVARENE
Grenoble Image Parole Signal Automatique, GIPSA-Lab, CNRS UMR5216
Université Joseph Fourier
Grenoble, France

MARTIAL MERMILLOD
Laboratoire de Psychologie Sociale et Cognitive, CNRS UMR6024
Université Blaise Pascal
Clermont-Ferrand, France

In this paper we present a neural network method for reconstructing a colour image from a mosaic a chromatic samples, arranged either in a regular or random manner. This method could be interpreted as a biological plausible method for the human visual system to reconstruct spatial and chromatic information from the random mosaic of cones in the retina.

1. Introduction

The human visual system has three kinds of cones for acquiring the colour in the visual environment in photopic (daylight vision) conditions. The cones have three spectral ranges of sensitivity and they are called L, M and S for their preferential sensitivity to Long, Middle and Short wavelength. Contrary to a colour image where three colours are sampled at each spatial position, in the human visual system, cones form a single mosaic where they are arranged randomly [1]. Consequently in the human retina only a single chromatic value is sampled for each spatial position at a time. This chromatic value corresponds to the type of cone L, M or S, which is present at the corresponding position in the retina. One can ask how the visual system is able to give us the sensation of colour and colour shading from a mosaic of chromatic samples. Moreover we can ask whether the random nature of the arrangement of these chromatic samples helps or diminishes our ability to perceive colour.

Another example of a mosaic of chromatic sampling is the digital camera. Actually, in most digital cameras today the acquisition of the colour image is done through a single sensor, which is covered by an array of chromatic filters. The so-called Colour Filter Array (CFA) gives the possibility for the sensor to discriminate between colours because the filter sensitivities cover three different ranges, usually Red, Green and Blue (RGB). Thus, in digital cameras we have to reconstruct three chromatic samples at each spatial position from an image with a single chromatic sample. This operation is called demosaicing [2]. In digital camera, the most used pattern is the Bayer CFA from the name of its inventor and consists of a regular pattern (Figure 4 (a)).

There is no real evidence that the human visual system reconstructs explicitly three colours information per spatial position such as the demosaicing process in cameras does [3]. Nevertheless, the spatial and chromatic information should be known at every position in the visual field to allow human to perceive shades of colours in natural scenes. Thus, we may suppose that even if the reconstruction of a colour image with three components does not occur in the visual system, this kind of representation of colour in an image can correspond to the information the human visual system needs to allow us to perceive colour. Thus, by extension, we suppose that the goal of the human visual system for colour perception could be compared to the ability to reconstruct three chromatic samples from a mosaic.

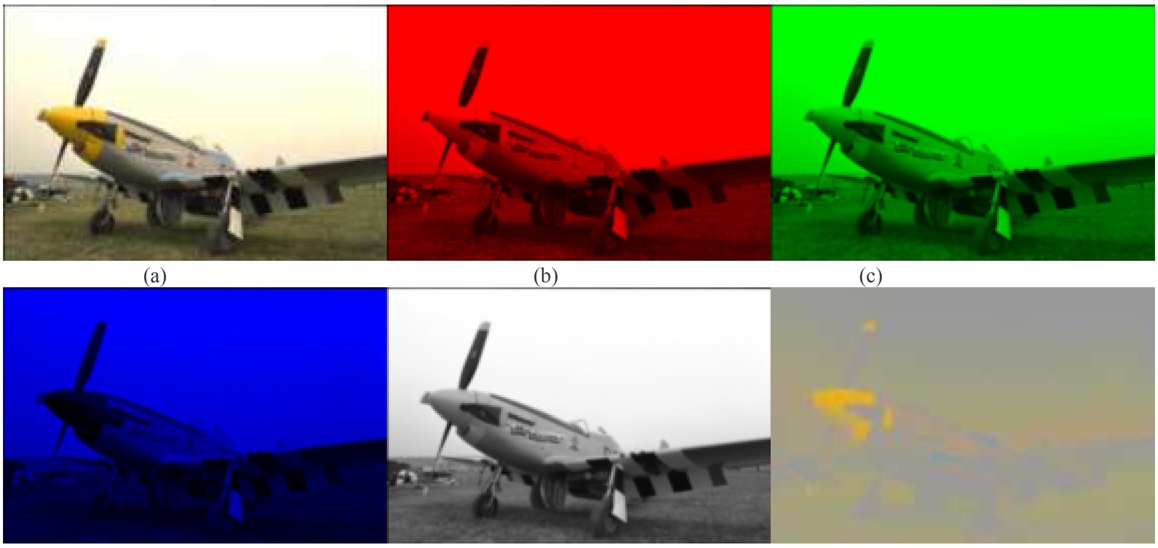
This article presents an approach where a neural network is used to reconstruct the missing colour information from a mosaic image. The method is applied on a regular and on a random arrangement of chromatic samples. The use of a neural network gives to the algorithm a biologic plausibility because the reconstruction is then given by a combination in a local neighbourhood of the existing pixels. This operation is plausible as a dendritic communication between neurons in the visual system.

2. Luminance-chrominance decomposition

The fact that the human visual system is sensitive to both achromatic spatial information and chromatic spatial information independently is well known since long ago [4]. This could be due to the way we perceived objects with their own colour automatically segregated from the shading of reflective light coming onto them.

There is no general consensus on the way the achromatic component is calculated from the chromatic component sampled by the retina. Actually, several decompositions were defined following the criterion used. The CIE recommends the use of the normalised function V(λ), which was measured on human subjects, to be the luminance visibility function.

The representation of a colour image into its achromatic and chromatic components is also interesting from a point of view of digital image processing. The following Figure 1 illustrates a decomposition of a colour image with its R, G and B values into luminance and chrominance components.



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

Figure 1: Illustration of the luminance - chrominance decomposition (a) RGB image (b) Red component (c) Green component (d) Blue component (e) luminance (f) chrominance.

We illustrate on Figure 1 the fact that the interpretation of the colour shade and intensity level of an image is easier in luminance and chrominance decomposition than in the RGB decomposition. As an example, consider the plane front-end. From Figure 1(e) and 1(f) it is clearly a yellow uniform colour with a shade of intensity from the top to the bottom. However, the same interpretation from Figure 1(b)-1(d) is difficult.

3. Model of spatio-chromatic sampling by a mosaic of chromatic samples

Let us define a colour image I with its three chromatic components, Red, Green and Blue, as follows:

$$I = \{C_i\}, i \in [R, G, B] \qquad (1)$$

We can formalize the sampling of a colour image through a mosaic by the following equation:

$$I_{mosaic} = \sum_i m_i(x,y)C_i(x,y) \tag{2}$$

Where m_i are the modulation functions, which take value 1 if the colour i is present at position (x,y) and 0 otherwise. Each modulation function m_i defines a submosaic of the chromatic samples of colour i. The modulation functions are specific to the arrangement of chromatic samples on the mosaic. But these functions can always be rewritten as a constant part plus a fluctuation part. Let the constant part called p_i be the mean value of the submosaic m_i. We may write:

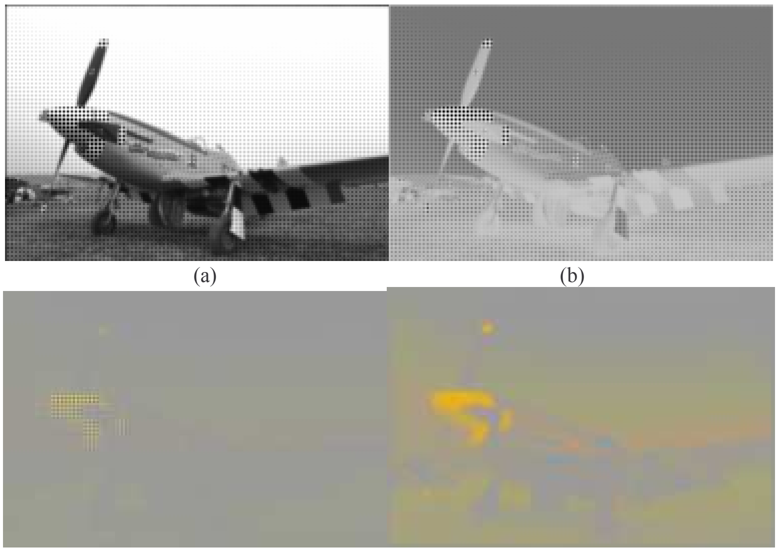
$$m_i(x,y) = p_i + \tilde{m}_i(x,y) \tag{3}$$

In that case, the constant part p_i corresponds to the proportion of each colour sample type in the mosaic. The fluctuation part $\tilde{m}_i$ takes the positive value $(1\text{-}p_i)$ in presence of the colour sample and negative values $-p_i$ elsewhere. With equation (3) we can rewrite equation (2) as follow:

$$I_{mosaic} = \underbrace{\sum_i p_i C_i(x,y)}_{Lum} + \underbrace{\sum_i \tilde{m}_i(x,y)C_i(x,y)}_{Chr} \tag{4}$$

Equation (4) shows that the mosaic image is in fact a sum of the luminance and the chrominance of the original image. The chrominance is actually subsampled and multiplexed in the mosaic image. The detail of the model can be found in Chaix et al. [5].

Figure 2 illustrates the composition of luminance and chrominance in mosaic images. In this simulation we used a mosaic image Figure 2(a) represented as a grey level image. We subtracted the luminance image defined in Figure 1(e) to the image in Figure 2(a) resulting in Figure 2(b). Then we demultiplexed the image on Figure 2 (b) by regrouping samples of the same colour sensitivity on their respective colour plane, either R, G or B. This allows representing that image in colour. Then we interpolate image on Figure 1 (c) resulting in image on Figure 2 (d). By comparing image on Figure 2 (d) and image on Figure 1(f) we can conclude that a mosaic image is a sum of the luminance image plus the subsampled and multiplexed chrominance image.

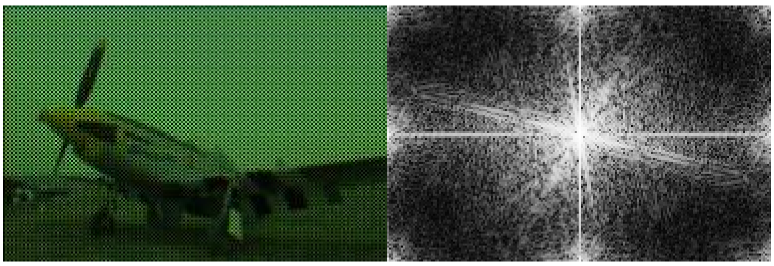


(a) (b)

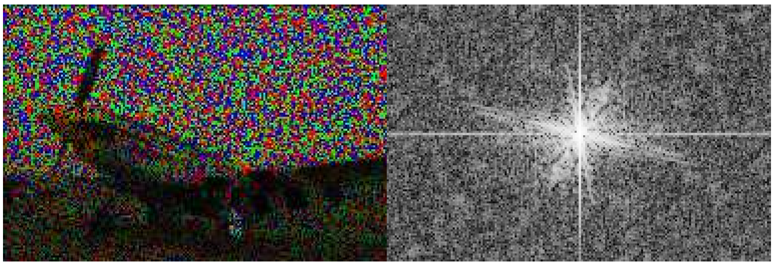
(c) (d)

Figure 2: Representation of luminance and chrominance in a mosaic. (a) The mosaic image displayed in grey levels. (b) Image in (a) where luminance in Figure 1(e) was removed. (c) demultiplexing of image (b) by isolating chromatic samples according to the mosaic arrangement. (d) Interpolation of image (c).

This decomposition does not depend on the arrangement of colour samples in the mosaic. In the case of a non-periodic arrangement, the demultiplexing is done the same way, following the position of each colour classes in the mosaic. Nevertheless, the interpolation of the chrominance could be more complex because the neighbourhood of colour changes from place to place in the mosaic.



(a) (b)



(c) (d)

Figure 3: Fourier representation of the mosaic. (a) The Bayer mosaic (b) The Fourier transform of the Bayer mosaic. (c) A mosaic with random arrangement of colour samples (d) The Fourier transform of the random mosaic.

The model of decomposition in luminance and chrominance allows interpreting the Fourier spectrum of the mosaic image [6]. In Figure 3 (b) we can identify nine different regions where the energy of the Fourier transform is localized. The region in the centre corresponds to the luminance signal whereas the regions in the border correspond to chrominance. In the case of a random arrangement of chromatic samples, the chrominance is no longer localized, but is rather spread over the whole Fourier domain as shown in Figure 3 (d). The localization of luminance and chrominance in the Fourier spectrum has been used to reconstruct colour image from the mosaic [6]. But this method does not apply in the case of random arrangement of chromatic samples in the mosaic. It should be found another method for the interpolation of chromatic samples in the case of random arrangement. Here we propose the use of a neural network.

4. Method

For simulating the process of sampling by a mosaic we used an image database[1] which comprises 24 colour images provided by Kodak. The images are originally of resolution 768x512 pixels. To reduce the amount of data we apply a subsampling of each individual image by a factor 4 in horizontal and vertical direction reducing the resolution to 192x128. To avoid aliasing, we apply the following low pass anti-aliasing filter before doing the subsampling.

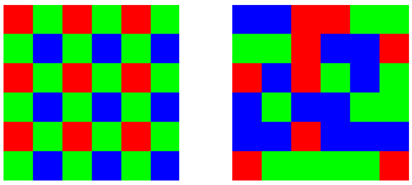
$$f = 2^{-16} g' g \text{ with } g' = [1 \quad 4 \quad 10 \quad 20 \quad 31 \quad 40 \quad 44 \quad 40 \quad 31 \quad 20 \quad 10 \quad 4 \quad 1] \quad (5)$$

The colour image database has three chromatic samples per spatial location. We remove chromatic samples accordingly to the arrangement in the mosaic to compose the input image. We then construct training and testing vector, using 12 images each.

[1] http://www.cipr.rpi.edu/resource/stills/kodak.html

To construct the vectors we use a neighbourhood of size 6x6 in the colour image as output vector of the neural network and the corresponding neighbourhood of size 12x12 in the mosaic image as input vector. Thus the neural network has the goal of reconstructing the three colours of 6x6 pixels from a neighbourhood of 12x12 pixels in the mosaic. We use a two layers neural network with an input layer and an output layer fully connected an initialized with random weights of mean zeros and variance 0.1. We use Lens (Light Efficient Network Simulator) [7] to perform the simulation. Once the network has converged, we use the output of the network given for the test set to reconstruct colour image. We can then compare the original colour image with the image reconstruct from the output of the neural network to test the quality of the reconstruction.

In a first experiment, we use either a Bayer or a random arrangement of size 6x6 as shown in the following Figure 4 (a) and (b) to construct the mosaic. It should be noted that we can not use a complete random arrangement for the chromatic samples because in this case the size of the input and output would be the same as the number of pixels in the image. To be able to train the network it is important that the position of chromatic samples was the same from a vector to another. Thus, we use a random pattern of size 6x6 that we tile along the image.



(a) (b)

Figure 4: (a) Representation of the Bayer mosaic on size 6x6 (b) Representation of the random mosaic of size 6x6 used for simulation.

In a second experiment, we use the neural network to estimate the luminance from the mosaic. Thus, the output vector is composed of a neighbourhood of 6x6 pixel of the luminance computed from the colour image.

5. Results

The result for the first experiment is given in Table 1 and Figure 5. The network converges quite quickly to a configuration that allows the reconstruction of colour. The following table (Table 1) shows the PSNR (Peak Signal to Noise Ratio) between the original colour image and the reconstructed image from the output of the neural network.

Table 1: PSNR between original image and image reconstruct from the mosaic by the neural network using either a Bayer or random mosaic.

	Bayer mosaic				Random mosaic			
Image	Red	Green	Blue	Total	Red	Green	Blue	Total
13	28.0667	26.9426	28.1920	**29.4156**	27.9650	26.9115	28.0575	29.2346
14	29.0981	30.4788	29.7122	**27.6237**	29.0862	30.5663	29.6630	27.5843
15	28.4966	29.5165	29.0924	**27.2336**	28.3348	29.3490	28.7703	27.1830
16	28.5257	29.7463	28.9645	**27.2510**	28.2301	29.5722	28.4295	27.0551
17	29.0636	30.4281	29.3598	**27.8071**	28.9417	30.2747	29.0963	27.8046
18	29.7367	29.4420	29.9980	**29.7883**	29.4121	29.2422	29.5339	29.4657
19	29.5661	30.4895	30.4456	**28.1836**	29.4003	30.3338	30.2847	28.0090
20	27.8741	27.8664	28.4647	**27.3612**	27.5003	27.6041	27.8171	27.1102
21	29.5636	30.0708	29.7752	28.9273	29.5528	29.9713	29.6641	**29.0716**
22	28.7614	29.4025	28.6200	28.3312	28.7111	29.6051	28.1625	**28.4934**
23	27.9490	28.1049	28.4094	**27.3959**	27.7253	27.9033	28.1576	27.1757
24	29.2061	30.3766	28.8397	**28.6059**	29.1300	30.3878	28.6826	28.5498

The following Figure 5 shows an example of the reconstruction of colour image from the mosaic.

(a) (b) (c)

Figure 5: Examples of reconstruction of a colour image (a) using the Bayer mosaic (b) or random mosaic (c).

We show that the neural network is able to learn how to reconstruct the colours from the mosaic of chromatic samples in a case of regular or random arrangement of chromatic samples. The result is not perfect and it remains some error in the reconstruction such as a blurring and some false colour. However we have not tested it extensively yet and it is possible that the structure of the network or the algorithm for convergence could be improved. We also show that the global level of luminosity is not exactly reconstructed. This is certainly due to the bias used in the neural network. Finally, although the Table 1 shows that the PSNR is often higher in the Bayer case than in the random case, the images show that the random case could be better when a high frequency pattern is under consideration. This can be shown for example on the window image of Figure 5, where the false colours are reduced in the random case. This could be due to the property of the random arrangement of chromatic samples to reduce the coherence of the error of reconstruction and hence to decrease its visibility. The result for the second experiment is shown in Table 2.

Table 2: Result for the second experiment

Image	Bayer	Random
13	**29.3712**	29.2042
14	30.9713	**31.1051**
15	**29.5052**	29.3472
16	**29.6361**	29.2351
17	**29.9640**	29.7346
18	**30.6322**	30.3208
19	**31.1623**	30.9672
20	**29.2731**	28.8540
21	**30.4569**	30.4035
22	**29.7240**	29.5396
23	**29.1436**	29.0356
24	29.8412	**29.9614**

Table 2 shows that the luminance is better reconstructed than R, G and B channel in term of objective PSNR criterion. This can be a new argument why the human visual system preferentially uses a luminance and chrominance approach than direct colour processing.

We have also tried to train a neural network with the chrominance information to see if the network was able to interpolate the chrominance. Unfortunately the network does not converge in that case. Thus, we are not able to show the colour results for the case of luminance-chrominance decomposition. The reason why the network diverge is not known, it may be due to the special statistics of the chrominance information of natural scene.

6. Discussion

In their paper, Doi et al. [8] show that the receptive field of ganglion cells at the output of the retina could be explained as an independent component analysis of natural scene when sampled by the random arrangement of chromatic cones in the retina. Here we show that the spatial and chromatic information could even be reconstructed from the mosaic with a good accuracy. We also show that the approach consisting of estimating luminance gives better results, although we are not able to reconstruct the chrominance by that way.

In this paper we used a supervised implementation of the neural network, given the input and the output to reach the convergence of the network. In the visual system if this kind of reconstruction happens, it remains to explain how the network would converge without having recourse to a supervised method for learning.

References

1. A. Roorda and D.R. Williams *Nature* **397** (11), pp.520-522 (1999).
2. B.K. Gunturk, J. Glotzbach, Y. Altunbasak, R.W. Schafer, and R.M. Mersereau, *IEEE Signal Processing Magazine*, **22** (1), pp.44-54, (2005).
3. F.A.A. Kingdom, and K.T. Mullen, *Spatial Vision*, **9**, pp.191-219 (1995).
4. E. Hering, «Zur Lehre vom Lichtsinn», Wien 1878.
5. B. Chaix de Lavarène, D. Alleysson and J. Hérault, *Computer Vision and Image Understanding*, **107**, (1), pp.3-13 (2007).
6. D. Alleysson, S. Süsstrunk and J. Hérault, *IEEE Trans. Image Processing*, **14** (4), pp. 439 - 449, (2005).
7. D. Rohde, Lens (Light Efficient Network Simulator) http://tedlab.mit.edu/~dr/Lens/
8. E. Doi, T. Inui, T-W Lee, T. Wachtler, and T.J. Sejnowski TJ *Neural Computation* **15**(2) pp.397-417, (2003).