RGBN Multispectral Images: A Novel Color Restoration Approach

Cristhian Aguilera[1], Xavier Soria[1], Angel D. Sappa[1,2](✉), and Ricardo Toledo[1]

[1] Computer Vision Center, Edifici O, Campus UAB,
08193 Bellaterra, Barcelona, Spain
{caguilera,xsoria,asappa,ricard}@cvc.uab.es
[2] Facultad de Ingeniería en Electricidad y Computación, CIDIS,
Escuela Superior Politécnica del Litoral, ESPOL, Campus Gustavo Galindo,
09-01-5863 Guayaquil, Ecuador

Abstract. This paper describes a color restoration technique used to remove NIR information from single sensor cameras where color and near-infrared images are simultaneously acquired—referred to in the literature as RGBN images. The proposed approach is based on a neural network architecture that learns the NIR information contained in the RGBN images. The proposed approach is evaluated on real images obtained by using a pair of RGBN cameras. Additionally, qualitative comparisons with a naïve color correction technique based on mean square error minimization are provided.

Keywords: Multispectral imaging · Free sensor model · Neural network

1 Introduction

The computer vision field has been growing considerably during the last decades. It is now ubiquitous in our every day life for a number of applications. Some of these applications rely for instance on image processing techniques such as image segmentation or object classification. These image processing techniques are based on photometric information obtained from the scene, such as color or relative difference between the color of the different regions. However, there is almost no difference between two different materials with the same color under a given light source; in other words, it is not possible to discriminate objects' material if they have exactly the same color. A possible solution for such a challenging problem can be found if we consider information from the electromagnetic spectrum beyond the range that the human visual system is sensitive to (400–700 nm), were visually similar samples may exhibit very different characteristics (e.g., [1–4]).

Such a property has been largely exploited in remote sensing applications where different spectral bands are used to characterize elements, such as materials, vegetation, water pollution, etc. (e.g., [5,6]). Among all the different spectral

F. De la Prieta et al. (eds.), *Trends in Cyber-Physical Multi-Agent Systems. The PAAMS Collection - 15th International Conference, PAAMS 2017*,
Advances in Intelligent Systems and Computing 619, DOI 10.1007/978-3-319-61578-3_15

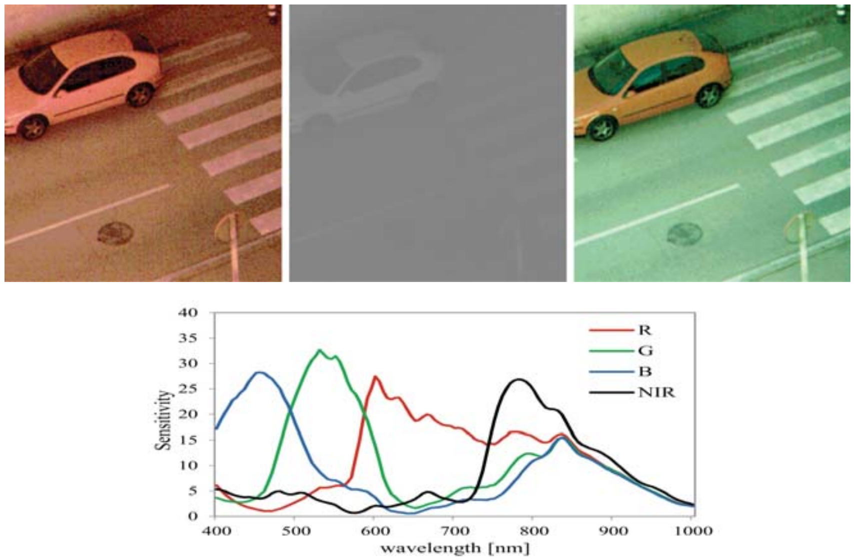

Fig. 1. ($top-left$) RGB image (obtained in the spectral band 400–1100 nm) with NIR information overlapped; ($top-middle$) NIR image (400–1100 nm) with RGB information overlapped. ($top-right$) RGB image (400–700 nm) used as ground truth; ($bottom$) Sensor's spectral sensitivity.

bands used in the computer vision beyond the visible spectrum, the near-infrared (NIR) is the most widely explored, since on the one hand it exhibits peculiar properties related with material energy absorption and reflectance; on the other hand sensors based on silicon (SiO_2) are sensitive to NIR up to 1100 nm, hence NIR camera are relative cheap in comparison with cameras working, with other technology, at other spectral bands. Since the NIR band is the closest in wavelength to the radiation detectable by the human eye, NIR images share several properties with visible images (see illustrations in Fig. 1). However, as mentioned above, surface reflection in the NIR band is material dependent. For instance, most dyes and pigments used for material colorization are somewhat transparent to NIR. This means that the difference in the NIR intensities is not only due to the particular color of the material, but also to the absorption and reflectance of dyes [1].

The absorption/reflectance properties mentioned above are used for instance in remote sensing applications for crop stress (water and nutrient stress being the most common) and weed/pest infestations. These applications are based on the fact that NIR is not absorbed by any pigments within a plant, it travels through most of the leaf and interacts with the spongy mesophyll cells. This interaction causes about half of the energy to be reflected and the other half to be transmitted through the leaf. In plants with turged and healthy mesophyll cell walls and in dense canopies, more NIR energy will be reflected and less transmitted. This cell wall/air space interaction within these cells causes healthy vegetation to look very bright in the NIR spectral band. In fact, much more NIR is reflected than visible. By monitoring the amount of NIR and visible energy reflected from the plant, it is possible to determine the health of the plant.

Trying to exploit such attractive properties (absorption/reflectance at NIR band and low cost technology) new cameras have been developed being able to work from the visible spectrum (400–700 nm) up to the NIR spectral band (700–1100 nm), providing the information in four independent channels—through this paper this technology will be referred to as RGBN cameras. Although interesting, the main problem with this technology lyes on the overlap between bands, as can be appreciated in Fig. 1, NIR sensor is sensible to RGB spectral bands, and part of NIR information goes also below 700 nm generating RGB images visually distorted (in general biased toward the red channel, see illustration in Fig. 1($top - left$)). The current work is focused on this problem (spectral band overlap) trying to restore color by removing NIR information. There are recent approaches in the literature, which are based on the sensor models (they will be reviewed in Sect. 2). In the current work a novel neural network based approach is proposed. It is trained by using RGBN and RGB images. The RGB images have been obtained using a IR cut-off-filter (IRCF), hence they are not affected by the NIR information. Promising results have been obtained independently of the lighting of the scene, which is a parameter needed in most of the sensor-based-model approaches.

The rest of the paper is organized as follows. In Sect. 2, both RGBN camera technology as well as previous work on RGBN color restoration from multispectral cameras are presented. Then, the proposed approach, which is based on the usage of a simple neural network, is introduced in Sect. 3. Experimental results are provided in Sect. 4. Finally, conclusions are given in Sect. 5.

2 Related Work

In the last years a new type of single sensor camera for simultaneous acquisitions of color and near-infrared images in the CMOS system has been proposed. Although it is an appealing technology, new problems have to be tackled in order to solve the band overlap problems discussed in the previous section. Additionally, the overlapping between the RGB and NIR channels could vary according to illumination, sensor characteristic and surface properties [7]. Hence, due to the complexity of the problem, some authors have proposed different solutions valid for controlled environments. For instance, [8] proposes a solution that works for controlled indoor scenarios, where the RGBN camera response is less influenced by NIR information. Other recent works have been proposed to solve this band overlapping problem (e.g., [9,10]), which is referred to in the literature to as color correction.

A color correction process is intended to obtain accurate color information. In [9] the authors propose the combined use of RGBN information for shadow removal and feature detection, although interesting results are presented

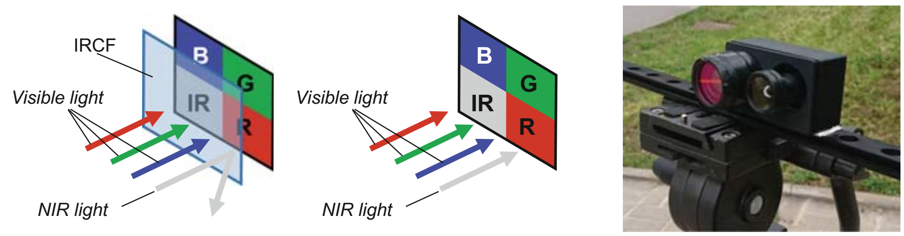

Fig. 2. Illustration of: (*left*) infrared filtered imaging system (IRCF: infrared cut-off filter); (*middle*) RGBN imaging system; (*right*) See3CAM-CU4 RGBN cameras used in the experiments. The infrared cut-off filter can be appreciated in the left side of the figure; it is used to generate the ground truth RGB images).

it should be noted that the proposed approach has been tested in a particular set of RGBN images obtained in a lab with fluorescent lamp. The obtained color corrected images were evaluated using a X-rite color checker chart, therefore in the visible spectrum there is not a considerable influence of sunlight; the tested scenario does not contains vegetation where the NIR influence plays an unknown and considerable role. This color correction in constrained scenario has been also considered in [11,12], although the authors mention sunlight presence in the indoor scenes tested. In these works the authors tackle the complex problem of restoring the three channels (R_{vis}, G_{vis} and B_{vis}), which are contaminated with an unknown quantity of NIR information. The proposed method is based on a spectral decomposition, it implies that the spectral response of each channel will correspond to the RGBN values. Each one of the channels contains a NIR and a visible spectrum part, which initially is formulated as follows: $NIR = NIR_{vis} + NIR_{nir}$, $R = R_{vis} + R_{nir}$ and so on, where the result is obtained as follows:

$$(\hat{R}_{vis}, \hat{G}_{vis}, \hat{B}_{vis})^T = \mathbf{M} \quad (R, G, B, N)^T, \tag{1}$$

where $\mathbf{M}$ is the decomposition matrix obtained by modeling the sensor sensitivity and band correlation; it is a scaling factor coefficient that relates the visible spectrum and NIR bands. Authors describe that the additional NIR information infected in the RGB channels maybe an unknown value. In other words, the spectral sensitivity curves presented in Fig. 1 depend on the sensor and are needed to solve Eq. 1. Note that the amount of NIR information will depend on both, the light in the scene and the type of material present on it. For instance, in outdoor scenarios, NIR information may change depending on the amount of vegetation, or materials with different absorption/reflectance properties (see Sect. 1).

Another image restoration technique has been recently proposed in [10]. In this case the visible image is obtained by subtracting at each visible channel a different amount of NIR information, according to coefficients previously computed. These coefficients are obtained from the sensor sensitivity curves (see Fig. 1). This particular formulation is only valid for a NIR wavelength range of

{650–819 nm}, since the camera used in that work is only sensible to the aforementioned range values. Although the results are quite good and the algorithm is efficient and fast to compute, its main drawback lies in the short wavelength validity. In [13], the authors propose a demosaicking[1] and color correction approach for improving the quality of acquired RGB and NIR images. The performance of this approach has been only evaluated using indoor images, when it is used in outdoor scenarios, the color correction does not work properly so that the obtained results are not like the natural colors.

3 Proposed Approach

We formulate the removal of the NIR information from the R_{RGBN}, G_{RGBN}, and B_{RGBN} channels of a RGBN camera as a regression problem. This regression problem is solved by using a neural network defined by two hidden layers with ten neurons each. The network is trained to learn a mapping function $\Omega : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ that maps a pixel color from an RGBN camera (see Fig. 2(*middle*)) to a RGB pixel value of the same scene, but obtained without NIR information (see Fig. 2($left$)).

The network model's input consists of a tuple $I = \{\mathtt{R}_{RGBN}, \mathtt{G}_{RGBN}, \mathtt{B}_{RGBN}, \mathtt{N}_{RGBN}\}$ that represent a color pixel value in a RGBN camera. The model has two hidden layers of ten neurons each and produces an output tuple $O = \{\mathtt{R}, \mathtt{G}, \mathtt{B}\}$ that contains the R,G,B values of the RGBN camera where the NIR information has been filtered.

We use a Smooth L1 loss function, defined as:

$$loss(x, y) = \frac{1}{n} \sum \begin{cases} 0.5 * (x_i - y_i)^2 & \text{if } |x_i - y_i| < 1 \\ |x_i - y_i| - 0.5 & \text{otherwise} \end{cases} \tag{2}$$

where, x is the set of all RGBN pixels values in the training set, and y the set of the corresponding pixels from the groundtruth.

The network has been trained using Stochastic Gradient Descent with the following parameters: learning rate 1e-5, momentum 0.17, weight decay 1e-5, and batchsize of 128. Figure 3 presents just an illustration of the proposed architecture.

4 Experimental Results

The proposed approach has been evaluated using two 4 megapixels single sensor RGBN cameras[2] rigidly attached in a common platform (see Fig. 2(*right*)). In one of the cameras an infrared cut-off filter is used; the obtained RGB images

[1] Demosaicking refers to obtaining the {R,G,B,NIR} components from a given pixel, where all the information is attached together is a single square array {B,G} in top and {IR,R} from the bottom, see an illustration of this pixel composition in Fig. 2.

[2] https://www.e-consystems.com/.

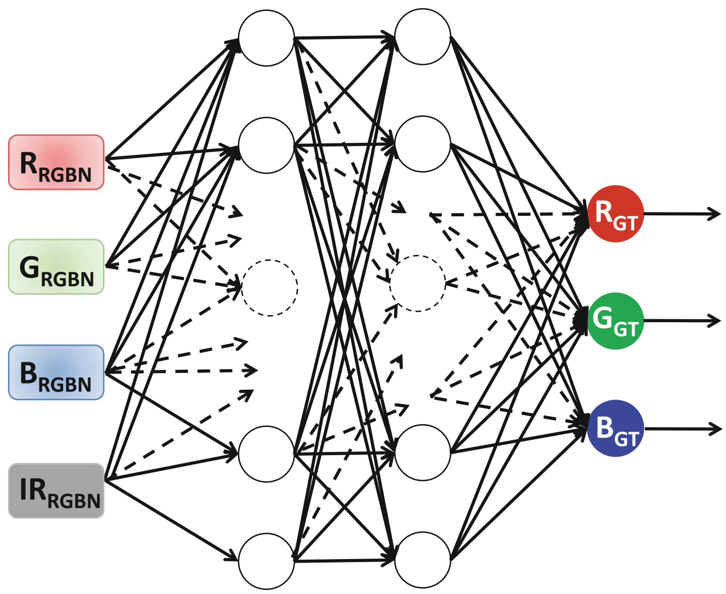

Fig. 3. Illustration of the network architecture used to learn the mapping function Ω.

are considered as ground truth. In order to qualitatively evaluate the obtained result a naïve approach based on minimizing the square error between the RGB and the corresponding RGBN images is provided; with this naïve approach a color correction matrix $\mathbf{M}_{CC}$ is obtained as follows:

$$E = \sum_{i=1}^{i=N} (RGB_{GT} - \mathbf{M}_{CC} \times RGB_{RGBN})^2, \tag{3}$$

where subindex GT corresponds to the ground truth values (elements from the camera with infrared cut-off-filter), subindex RGBN corresponds to the pixel values obtained with the RGBN camera, $\mathbf{M}_{CC}$ is the color correction matrix (a square matrix of 3×3 elements), and N represents all the pixels from all the images used to estimate (E). In order to do a fair comparison the same set of images (83 pairs of images of 256×256 pixels) used to train the network have been considered; in other words (E) is computed by considering $N = 83 \times 256 \times 256 = 5{,}439{,}488$ elements.

With the setup presented in Fig. 2(*right*) a set of 89 pairs of images has been obtained[3]. Note that although the cameras have been rigidly attached, trying to place their optical axis as parallel as possible, the obtained images need to be registered. This registration process is needed in order to guarantee a pixel-to-pixel correspondence. Differences due to camera disparity are neglected since cameras' optical axis are quite near in comparison to the depth of the

[3] http://www.cvc.uab.es/~asappa/Color_Correction_Dataset.rar.

(*a*) (*b*) (*c*) (*d*)

Fig. 4. Illustrations of: (*a*) Original RGBN image; (*b*) Result obtained from the MSE color correction; (*c*) Result from the network trained with RGB and NIR images; (*d*) RGB ground truth image obtained by using the infrared cut-off-filter.

objects in the scene (actually, the data set has been created having in mind this assumption, so that scenes containing objects far away from the camera have been considered). RGB and RGBN images have been registered using the Matlab Image Alignment Toolbox (IAT)[4] (in the RGBN images just the RGB channels are considered for the registration process). After images have been registered, they were cropped into images of 256 × 256 pixels to avoid problems with border pixels. This cropping area is centered in every registered image. From the 89 pairs data set 71 pairs have been used for both the network training and for computing the $\mathbf{M}_{CC}$; the remainder 18 pairs were used for validating the results. A qualitative validation (just 18 pairs) has been performed comparing the results from the neural network and those obtained by using the naïve color correction approach.

Figure 4 presents just three illustrations of the results obtained with the neural network proposed approach Fig. 4(*c*), which can be compared with the corresponding ground truth Fig. 4(*d*), as it can be appreciated, although the network architecture is quite simple, the results obtained with the proposed

[4] http://iatool.net/.

approach looks quite similar to those form the ground truth. On the contrary, the result obtained from the color correction matrix (see Fig. 4(b)) are not so similar to those from the ground truth. As a conclusion from these results we have identified two ways to improve them. The first one is related with the data set. A larger data set, with more color variability, obtained at different daylight time is required to learn a more accurate Ω mapping function by the network. The second improvement is related with the used architecture, other configurations should be evaluated (more hidden layers and more neurons per layer). In Fig. 4 just three results from the 17 pairs used for validation are presented, the whole set of images used for validation can be downloaded from: http://www.cvc.uab.es/~asappa/Results_for_Validation.rar.

5 Conclusions

This paper presents a novel approach to tackle the challenging problem of NIR information removal (referred to as color restoration) in single sensor RGBN cameras. On the contrary to previous approaches, which are based on a sensor model that work under constrained scenarios, the proposed approach is based on a neural network architecture that is able to learn a mapping function Ω that transform the 4D coupled information into a 3D representation. Experimental results with different outdoor scenarios have been tested showing the validity of the proposed approach. Additionally, comparisons with a naïve color correction based on minimizing the square error are provided.

Acknowledgments. This work has been partially supported by the Spanish Government under Project TIN2014-56919-C3-2-R and by the ESPOL projects: "Pattern recognition: case study on agriculture and aquaculture" (M1-DI-2015) and "Integrated system for emergency management using sensor networks and reactive signaling" (G4-DI-2014). Cristhian Aguilera has been supported by Universitat Autònoma de Barcelona. Xavier Soria would like to thank to Ecuador government under a scholarship contract 2015-AR3R7694.

References

1. Salamati, N., Fredembach, C., Süsstrunk, S.: Material classification using color and NIR images. In: Color and Imaging Conference, Society for Imaging Science and Technology, vol. 2009, pp. 216–222 (2009)
2. Ricaurte, P., Chilán, C., Aguilera-Carrasco, C.A., Vintimilla, B.X., Sappa, A.D.: Feature point descriptors: infrared and visible spectra. Sensors **14**, 3690–3701 (2014)
3. Barrera, F., Lumbreras, F., Sappa, A.D.: Evaluation of similarity functions in multimodal stereo. In: International Conference Image Analysis and Recognition, pp. 320–329, Springer (2012)
4. Mouats, T., Aouf, N., Sappa, A.D., Aguilera, C., Toledo, R.: Multispectral stereo odometry. IEEE Trans. Intell. Trans. Syst. **16**, 1210–1224 (2015)
5. Mountrakis, G., Im, J., Ogole, C.: Support vector machines in remote sensing: a review. ISPRS J. Photogrammetry Remote Sens. **66**, 247–259 (2011)

6. Adam, E., Mutanga, O., Rugege, D.: Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. Wetlands Ecol. Manag. **18**, 281–296 (2010)
7. Barnard, K., Cardei, V., Funt, B.: A comparison of computational color constancy algorithms. I: methodology and experiments with synthesized data. IEEE Trans. Image Process. **11**, 972–984 (2002)
8. Chen, Z., Zhu, N., Pacheco, S., Wang, X., Liang, R.: Single camera imaging system for color and near-infrared fluorescence image guided surgery. Biomed. Opt. Express **5**, 2791–2797 (2014)
9. Chen, Z., Wang, X., Liang, R.: RGB-NIR multispectral camera. Opt. Express **22**, 4985–4994 (2014)
10. Martinello, M., Wajs, A., Quan, S., Lee, H., Lim, C., Woo, T., Lee, W., Kim, S.S., Lee, D.: Dual aperture photography: image and depth from a mobile camera. In: 2015 IEEE International Conference on Computational Photography (ICCP), pp. 1–10 (2015)
11. Park, C.H., Oh, H.M., Kang, M.G.: Color restoration for infrared cutoff filter removed RGBN multispectral filter array image sensor. In: VISAPP, (1), pp. 30–37 (2015)
12. Park, C., Kang, M.G.: Color restoration of rgbn multispectral filter array sensor images based on spectral decomposition. Sensors **16**, 719 (2016)
13. Teranaka, H., Monno, Y., Tanaka, M., Ok, M.: Single-sensor RGB and NIR image acquisition: toward optimal performance by taking account of CFA pattern, demosaicking, and color correction. Electronic Imaging 2016, pp. 1–6 (2016)