Electronic Letters on Computer Vision and Image Analysis 19(2):5-8, 2020

Processing and Representation of Multispectral Images Using Deep Learning Techniques

Patricia L. Suarez

* *Faculty of Electrical and Computer Engineering, ESPOL Polythecnic University, Guayaquil, Ecuador*
Thesis advisor: Angel Sappa, Co-advisor: Boris Vintimilla
Date and location of PhD thesis defense: 17 April 2020, ESPOL Polythecnic University

Received 10 June 2020; Accepted 26 October 2020

Abstract

This thesis has implemented innovative techniques in the field of computer vision using visible and near-infrared spectrum images, applying deep learning through convolutional networks, especially GANs' architectures, and also includes meta-learning techniques to tackle the proposed problems. In this research, with this type of convolutional networks, different supervised and unsupervised techniques have been created to solve challenging problems, like detect the similarity of patches of different spectra (visible-infrared), colorized images of the near-infrared spectrum, estimation of vegetation index (NDVI) and the haze removal present on RGB images using NIR images. For all these techniques different variants of the GAN's networks, such as standard, conditional, stacked, and cyclic have been used. Also, a metric-based meta-learning approach has been implemented. It should be mentioned that together with the implementation of adversarial network models, the use of multiple loss functions has been proposed to improve the generalization and increase the effectiveness of the models. The experiments were performed with paired and unpaired images, given the different supervised and unsupervised architectures implemented, respectively. The experimental results obtained in each of the approaches implemented in the doctoral work compared with the techniques of the state of the art were shown to be more effective.

Key Words: Convolutional Neural Networks, Generative Adversarial Network, Infrared Imagery colorization, Haze, Normalized Difference Vegetation Index, Stacked Generative Adversarial Network.

1 Introduction

Ultimately, in the field of computer vision, techniques based on the use of several spectra, not only the visible spectrum, are being proposed to address problems of detection, recognition, composition of materials, surface characteristics, among others. For example, with the use of cross-spectral images (visible and near-infrared), additional information is obtained for each spectrum that can be used to improve the different existing visualization techniques. With these images, great advantages can be obtained in recognition and detection tasks. To acquire this type of cross-spectral data set, multiple cameras are often used, which requires alignment or

Correspondence to: <plsuarez@espol.edu.ec>

Recommended for acceptance by <Christian Aguilera and Arash Akbarinia>
https://doi.org/10.5565/rev/elcvia.1246
ELCVIA ISSN:1577-5097
Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

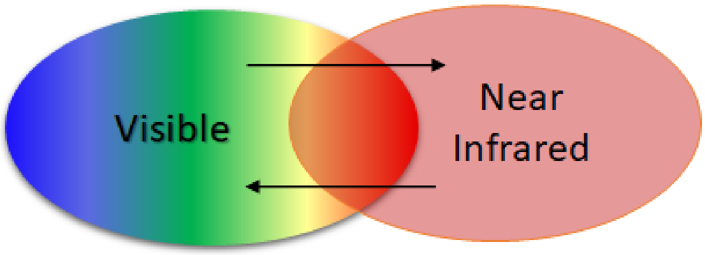

Figure 1: The combined use of different bands of the electromagnetic spectrum:

estimation of the disparity of the images. Increasingly, multi-camera cross-spectrum systems are integrated into active RGBD devices (for example, RGB-NIR cameras in Kinect and iPhone X).

Generally, there are some limitations with the lighting conditions, material texture or temperature that considerably affect the performance of computer vision techniques when only visible spectrum images are used. These limitations can be overcomed using cross-spectral images, to tackle the design of computer vision problems more effectively.

The simultaneous use of images from different spectra can be helpful to improve the performance of many computer vision tasks. The core idea behind the usage of cross-spectral approaches is to take advantage of the strengths of each spectral band providing a richer representation of a scene, which cannot be obtained with just images from one spectral band. See Fig. 1.

There are some approaches designed to use images from different spectral bands to exploit all features inherent to those bands other than visible and in that way improve the performance of the techniques. Some of these techniques can be applied in the following ways; in surveillance systems, using the dense depth information on images from the thermal and visible spectrum to determine the differences between them and obtain adequate estimates for object localization or environment navigation [5].

2 Objectives

This thesis focuses on the exploration of the use of information from the different bands of the electromagnetic spectrum in such a way that it can be exploited to solve some problems existing in the field of computer vision, for which different deep learning architectures have been designed to be used with multispectral images. The approaches addressed in the thesis are detailed below:

1. Cross-Spectral Image Similarity: Determine similarity with a performance similar or better than methods based on visible spectra [1].

2. NIR Image Colorization: Implement a novel colorization process using NIR images [3].

3. NDVI Vegetation Index Estimation: Support the analysis of the health of the vegetation, using only sensors sensitive to the visible spectrum. [7].

4. Image Dehazing: Improve the quality of the image using near-infrared spectra images [6].

3 Proposed Approaches

To tackle the problem of determining the similarity of images of different spectra, two approaches have been proposed, the first one using CNN and the second one based on meta-learning; from the dataset [2], the categories: *country, indoor, oldbuilding and urban* have been selected. These images are the most affected in lighting conditions and textures, which directly affect the complexity of the process of establishing their similarity through the detection of characteristic points and, therefore, are the most challenging scenarios for the training process. The first approach learn the similarity between cross-spectral image patches with a 2 channel convolutional neural network (CNN) model. The second approach is a technique based on meta-learning, which proposes an 8-$shot$ 1-way meta-learning metric based network model. The results of the experiments of the two approaches implemented have been compared to each other and also with another similar technique presented in [1].

In the case of the NIR image colorization problem, four approaches based on adversarial generative networks have been implemented, the first approach has been implemented using an architecture with a single learning level, FLAT, which obtained the first satisfactory results of NIR colorization has subsequently been implemented, a GAN colorization architecture based on three learning levels (triplet), one for each color channel in the RGB space, with which better metrics were obtained in the colorization results, Continued with the experiments, a third approach has been implemented, using a conditional GAN architecture applied to the triplet, together with the application of the multiple loss function to generalize the faster model and avoiding local maximums. Finally, a stacked GAN model has been implemented together with a new loss function, it has been possible to obtain the best qualitative and quantitative results for the colorization of the NIR images comparing them with the previous implemented approaches and with other state of the art colorization approaches.

To address the problem of NDVI vegetation index generation, supervised GAN and unsupervised cyclical networks have been implemented. For the first case, a conditioned triplet architecture has been used, the second approach has been based on the use of a conditioned GAN network of a single learning level with multiple loss function, the third approach has been implemented in a non-schematic supervised, using cyclical GAN networks, to obtain a synthetic NIR with which to calculate the NDVI index and finally the fourth approach, using a cyclic GAN network to obtain the vegetation index from only the red channel of an image of the RGB color space , which has been the approach that has obtained better qualitative and quantitative results comparing them with the state of the art.

Finally, to solve the problem of removing haze from images, two approaches have been implemented, the first a conditioned GAN network that uses only RGB color space images along with multiple loss functions. The second approach is a GAN conditioned on a NIR image that is used to enhance the details of the images. You can read all approaches in detail in the *doctoral thesis*, which can be accesed in *.

References

[1] Cristhian A Aguilera, and Francisco J Aguilera, and Angel D Sappa, and Cristhian Aguilera, and Ricardo Toledo, "Learning cross-spectral similarity measures with deep convolutional neural networks", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 1-9, 2016

[2] Brown, Matthew and Süsstrunk, Sabine,"Multi-spectral SIFT for scene category recognition", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 177-184, 2011

[3] Limmer, Matthias and Lensch, Hendrik PA, "Infrared colorization using deep convolutional neural networks", *15th IEEE International Conference on Machine Learning and Applications*, pages 61-68, 2016

[4] Huang, Xun and Li, Yixuan and Poursaeed, Omid and Hopcroft, John and Belongie, Serge, "Stacked generative adversarial networks", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5077-5086, 2017

* http://refbase.cidis.espol.edu.ec/files/patriciasuarezriofrio/2020/122_PatriciaSuarezRiofrio.2020.pdf

[5] Pinggera, Peter and Breckon, Toby and Bischof, Horst, ”On cross-spectral stereo matching using dense gradient features”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2012

[6] Wang, Jin-Bao and He, Ning and Zhang, Lu-Lu and Lu, Ke,”Single image dehazing with a physical model and dark channel prior”, *Neurocomputing*, volume 149, pages 718-728, 2015, publisher=Elsevier

[7] Kamilaris, Andreas and Prenafeta-Boldú, Francesc X, ”Deep learning in agriculture: A survey”, *Computers and Electronics in Agriculture*, volume 147, pages 70-90, 2018, publisher Elsevier