Adaptive Harris Corner Detector Evaluated with Cross-Spectral Images

Patricia L. Suárez[1(✉)], Angel D. Sappa[1,2], and Boris X. Vintimilla[1]

[1] Facultad de Ingeniería en Electricidad y Computacíon, CIDIS, Escuela Superior Politécnica del Litoral, ESPOL, Campus Gustavo Galindo, 09-01-5863 Guayaquil, Ecuador
{plsuarez,asappa,boris.vintimilla}@espol.edu.ec

[2] Computer Vision Center, Edifici O, Campus UAB, Bellaterra, 08193 Barcelona, Spain

Abstract. This paper proposes a novel approach to use cross-spectral images to achieve a better performance with the proposed Adaptive Harris corner detector comparing its obtained results with those achieved with images of the visible spectra. The images of urban, field, old-building and country category were used for the experiments, given the variety of the textures present in these images, with which the complexity of the proposal is much more challenging for its verification. It is a new scope, which means improving the detection of characteristic points using cross-spectral images (NIR, G, B) and applying pruning techniques, the combination of channels for this fusion is the one that generates the largest variance based on the intensity of the merged pixels, therefore, it is that which maximizes the entropy in the resulting Cross-spectral images.

Harris is one of the most widely used corner detection algorithm, so any improvement in its efficiency is an important contribution in the field of computer vision. The experiments conclude that the inclusion of a (NIR) channel in the image as a result of the combination of the spectra, greatly improves the corner detection due to better entropy of the resulting image after the fusion, Therefore the fusion process applied to the images improves the results obtained in subsequent processes such as identification of objects or patterns, classification and/or segmentation.

Keywords: Near Infrared · Cross-spectral · Visible spectra · Pixel Fusion · Pruning

1 Introduction

Computer vision tackles problems related with object detection and recognition, texture classification, action recognition, segmentation, tracking, data retrieval, image alignment, just to mention a few. In general, computer vision solutions are based on representing the given image using some global or local image properties [1], and then comparing them using some similarity measure [2]. Additionally,

Á. Rocha and T. Guarda (eds.), *Proceedings of the International Conference on Information Technology & Systems (ICITS 2018)*, Advances in Intelligent Systems and Computing 721,
https://doi.org/10.1007/978-3-319-73450-7_69

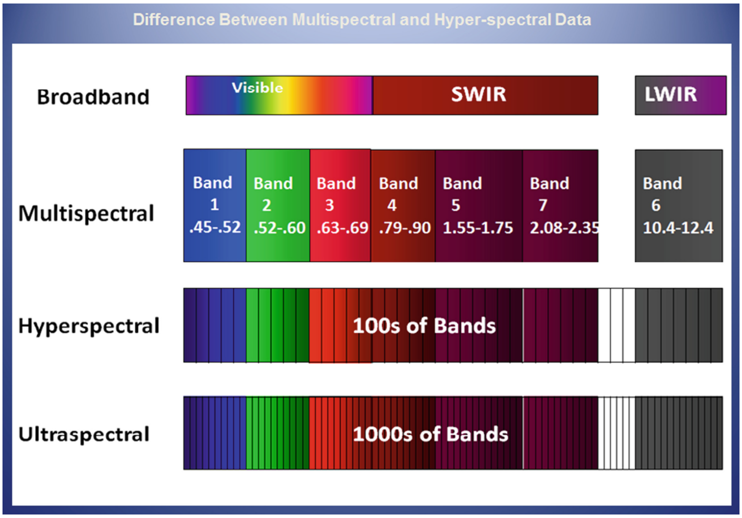

Fig. 1. Spectrum Band

computer Vision plays a central role in human perception and interpretation of the world. Our eyes and brain can quickly provide detailed information about any event that is happening around us, leading to an appropriate choice of action or response. The importance of human visual perception is also evident when one considers that the vision processing consumes a proportionately large part of the function of the human brain.

The most significant recent advances in remote sensing has been the development of Multi-spectral sensors and specialized programs to analyze the data of the resulting images. Just fifteen years ago remote sensing experts have had access to the Multi-spectral imaging and specialized tools to take advantage of the information they provide such images [3]. In the last decade the analysis of multi- spectral imaging has matured and has become one of the fastest growing technologies in the field of remote sensing. The term "Multi" in Multi-spectral means "many" and refers to the large number of wavelength bands that constitute them. Multi-spectral images provide a wide spectral information to appreciate characteristics that can not be seen in other spectral band. See spectrum band in Fig. 1. As the volume of hyper-spectral data for planetary exploration increases, efficient yet accurate algorithms are decisive for their analysis. The capability of spectral unmixing for analyzing hyper-spectral images from Mars is now under investigation [4]. Hyper-spectral monitoring of large areas (more than 10 km^2) can be achieved via the use of a system employing spectrometers and CMOS cameras. A robust and efficient algorithm for automatically combining multiple, overlapping images of a scene to form a single composition (i.e., for the estimation of the point-to-point mapping between views), which uses only the information contained within the images themselves [5].

Advances in remote sensing technologies are increasingly becoming more useful for resource, ecosystem and agricultural management applications to the extent that these techniques can now also be applied for monitoring of soil contamination and human health risk assessment. While, extensive previous studies have shown that Visible and Near Infrared Spectroscopy (VNIRS) in the spectral range 400–2500 nm can be used to quantify various soil constituents simultaneously, the direct determination of metal concentrations by remote sensing and reflectance spectroscopy is not as well examined as other soil parameters [6].

Multi-spectral images provide the potential for more accurate and detailed information extraction and that is not possible with any other type of remote sensing data. Remote sensing applications are usually applied in projects that generally have one of the following objectives: Detection of targets (objects, tumors, people, others), Mapping Materials, Tracing, Classification, Segmentation, Mapping the surface properties in material identification.

Digital image fusion has been used in research field since the late nineties at the leading edge of available technology. It formed a rapidly developing area of research in remote sensing [7]. Recently, cross-spectral based approaches are obtaining remarkable results in computer vision applications, as well as in a large number of fields. For instance, using this king of images has been recently presented with success working on images in the mono-spectral or in cross-spectral domain [8–10].

In order to prepare the samples of the fused images, the bands of the images of the visible spectrum (R-red, V-green, B-blue) are separated and merged with the near-infrared image. It is done the fusion of images with the combination of channels (NIR, G, B) with which the best entropy is obtained. See Fusion Process in Fig. 2. Thus it is possible to start the experiments and to use the merged images for the detection of corners by means of the Adaptive Harris algorithm and in such a way to demonstrate that there is an improvement of the results of the detection of the characteristics of the images, a similar work using cross-spectral imagery was performed to improve edge detection using morphological operations [11].

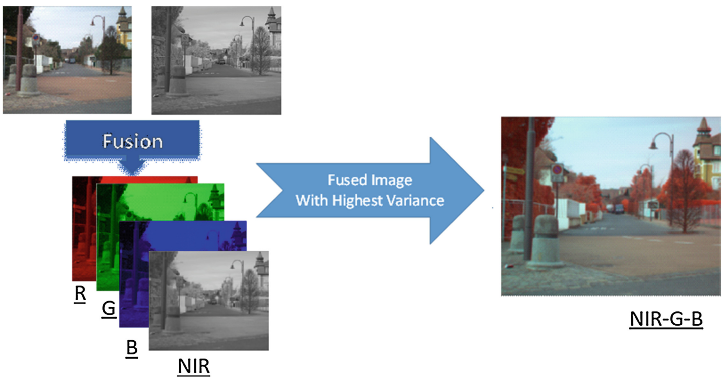

Fig. 2. Cross-Spectral fusion process

The corners can be defined as the points with the lower similarity in all directions, this can be measured by taking the sum of squares differences (SSD). The Harris algorithm works by calculating a response function across all pixels of the images. See Corner representation in Fig. 3. After that, those who exceed the threshold of its local maximum are recognized as corners and they are retained.

In computer vision, it is necessary to establish matching points between different images, this allows us extract information and to take action on them. When we talk about matching points we refer generally to the characteristics of the scene that we need to recognize easily and uni. What are the most important features, to name a few are: borders, regions and corners. But of all, the corners are the most important feature because being the intersection of two edges, they represent a point where the direction changes, therefore, the gradient of the image has a high variation, which is used to detect it.

According to [12] the traditional Harris Algorithm considers a grayscale image I. We are going to sweep a window $\boldsymbol{w(x,y)}$ (with displacements $\boldsymbol{u}$ in the $\boldsymbol{x}$ direction and $\boldsymbol{v}$ in the right direction) I and will calculate the variation of intensity:

$$\mathbb{E}(u,v) = \sum_{x,y} w(x,y)[I(x+u,y+v) - I(x,y)^{(2)})], \tag{1}$$

where: $\boldsymbol{w(x,y)}$ is the window to the position $\boldsymbol{(x,y)}$, $\boldsymbol{I(x,y)}$ is the intensity at $\boldsymbol{(x,y)}$ and $\boldsymbol{I(x+u,\ y+v)}$ is the intensity at the moved window $\boldsymbol{(x+u,\ y+v).}$

Since we are looking for windows with corners, we are looking for windows with a large variation in intensity. Hence, we have to maximize the equation above, so using Taylor expansion, expanding the equation and canceling properly, it can be expressed in a matrix form as:

$$\mathbb{E}(u,v) \approx [u,v]\left(\sum_{x,y} w(x,y)\begin{bmatrix} I_x^2 & I_x y \\ I_x y & I_y^2 \end{bmatrix}\right)\begin{bmatrix} u \\ v \end{bmatrix} \tag{2}$$

then let's denote:

$$M = \left(\sum_{x,y} w(x,y)\begin{bmatrix} I_x^2 & I_x y \\ I_x y & I_y^2 \end{bmatrix}\right) \tag{3}$$

So, the equation now is:

$$\mathbb{E}(u,v) \approx [u,v]M\begin{bmatrix} u \\ v \end{bmatrix} \tag{4}$$

Harris algorithm works by calculating a response function (RF) through all pixels of the image, after which, those who exceed the threshold, which is also known as a local maximum are retained as corners. Being (I) a 2D image grayscale. Measure response of a corner:

$$R = \det(M) - k(trace(M))^2 \tag{5}$$

where: $\det(M) = \lambda_1, \lambda_2$, $trace(M) = \lambda_1 + \lambda_2$, λ_1 and λ_2 are the *eigenvalues* of M and k are a *empiric* constant between 0.04 and 0.06.

R depends only on the eigenvalues of M. Therefore defined as follows: If R is large corresponds to a corner, if R is negative with large magnitude it corresponds to an edge and if $|R|$ is small corresponds to a flat region. Harris detection algorithm finds the points with the greatest values in the response function corners (R) and working with a threshold. Points are taken with a local maximum R. The quality of the detected corners depend on the threshold used to discern them. A quite high threshold will detect only very strong corners, while a too low threshold will detect many false corners, which are originated by noisy points, this can be computationally expensive, which is why new variants have been developed that allow to reduce the amount of valid information to be processed in the process of detection of corners.

In this context, the current paper tackles a more robust feature extractor using fused cross-spectral images and use a variation of Harris, it is a novel low complexity pruning technique that removes the non-corners using simple approximations of the complex Harris corner measure to create a small corner candidate set [13], that allow to obtain more efficiently the principal key points that correspond to a corner, and this make relevant the improvement to facilitate process like object detection, image classification, panoramic scenes creation and 3D generated image reconstruction. The rest of the paper is organized as follows. Section 2 describes the most recent work on feature extraction based on remote sensing for several estimations. Section 3 presents the Adaptive Harry Corner Detection approach with Cross-Spectral images proposed. Section 4 depicts the experimental results and finally, conclusion are presented in Sect. 5.

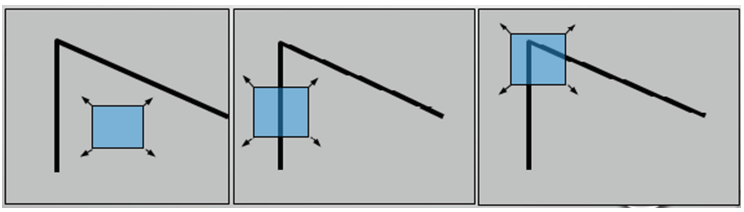

Fig. 3. Corner representation

2 Related Work

Remote sensing has become a major source of land use information to cover a range of spatial scales and temporal scales. Recently, Some papers have focused on the improvement of the classification process; others on the use of well-known classification methods in particular types of remote sensing application. Classification is regarded as a fundamental process in remote sensing, which lies at the heart of the transformation from satellite image to usable geographic product [14].

Traditional classification techniques assigns each pixel to a single class, using remote sensing images. In particular areas with high spatial resolutions, they are commonly dominated by mixed pixels that contain more than one class in the soil. For the sub-pixel information is needed to use techniques and soft sorting algorithms in Multi-spectral domain order to obtain the fraction of each class in a pixel. There are a proposed work based in a Linear spectral unmixing which is a popular tool in remotely sensed hyper-spectral data interpretation. It aims at estimating the fractional abundances of pure spectral signatures (also called as endmembers) in each mixed pixel collected by an imaging spectrometer [15]. Another approach, is the evaluation of an image that depend upon the purpose for which the image was obtained and the manner in which the image is to be examined. Where the goal is extraction of information and where the image is to be processed prior to viewing, the information content of the image is the only true evaluation criterion. Under these conditions, the improvement achieved by processing can be evaluated by comparing the ability of the human observer to extract information from the image before and after processing. The extent to which the processing approaches the optimum can be evaluated by determining the fraction of the total information content of the image which can be visually extracted after processing [16]. Another technique it is based on a set of high-resolution remote sensing images covering multiple spatial features, they proposed an classification based on unsupervised technique including pixel-wise and sub-pixel-wise methods to detect possible built-up areas from remote sensing images. The motivation behind is that the frequently recurring appearance patterns or repeated textures corresponding to common objects of interest in the input image data set can help us distinguish built-up areas from other features [17]. Another paper proposes a froth image segmentation method combining image classification and image segmentation. In the method, an improved Harris corner detection algorithm is applied to classify froth images first. Then, for each class, the images are segmented by automatically choosing the corresponding parameters for identifying bubble edge points through extracting the minimal local gray value. Finally, on the basis of the edge points, the bubbles are delineated by using a number of post-processing functions. Compared with the widely used Watershed algorithm and others for a number of lead zinc froth images in a flotation plant, the new method (algorithm) can alleviate the over-segmentation problem effectively.

Other approach proposed a multi-target tracking algorithm based on a particle filter framework that exploits a sparse distributed shape model to handle partial occlusions. The state vector is composed by a set of points of interest (i.e. corners) and it enables to jointly describe position and shape of the target. An efficient importance sampling strategy is developed to limit the number of

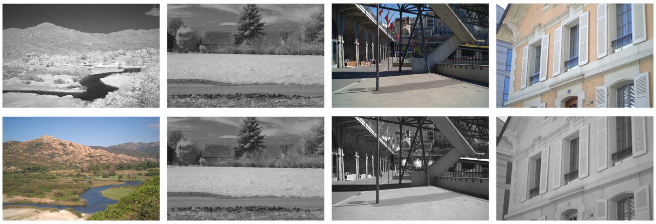

Fig. 4. Pairs of images (1024 × 680 pixels) from [18]; *country* category (the first column), *field* category (the second column); *urban* category (the third column) and *old-building* category (the fourth column); (*top*) NIR images; (*bottom*) RBG images.

used particles and it is based on multiple Kanade-Lucas-Tomasi (KLT) feature trackers used to estimate local motion [19]. The usage of cross-spectral information, although interesting and appealing, implies new challenging and difficult problems that need to be tackled and efficiently solved. For instance, different works have been recently proposed for describing and matching feature points in cross-spectral domains based on classical approaches. In the current work we propose the use of Cross-Spectral images to improve the image entropy and use and adaptive pruning variation before applied traditional Harris algorithm to reduce the computational cost, improve the accuracy, because significantly reduces the selection and evaluation effort for the presence of corners to only corner like regions.

3 Proposed Approach

This section presents the approach proposed for improve the corners detection process, based on the usage of a Cross-Spectral images and an enhanced pruning technique before to apply the conventional Harris algorithm, we propose to use them to achieve a better performance and accuracy in the extraction of features and at the same time reduce the computational cost of the corner detection process. This can be done because this pruning technique uses a new threshold model where product of vertical and horizontal difference in pixel intensities is used and the candidates with low CR (corner response) values are pruned away.

An adaptive corner response (CR) approach is defined as:

$$CR = \left(\left| I_x \cdot I_y \right| \right) \tag{6}$$

The Harris detector compute the corner measure on every image pixel and the obvious non-corners are removed by applying a threshold on the corner measure. The corner response of pixels that are close to a good corner are also typically high and hence, a minimum distance is enforced between good corners using non-maximal suppression (NMS). Finally all the corners are sorted and only the top few corners are selected for further processing. The high computational load

Fig. 5. Cross-Spectral image, resulting from the fusion process

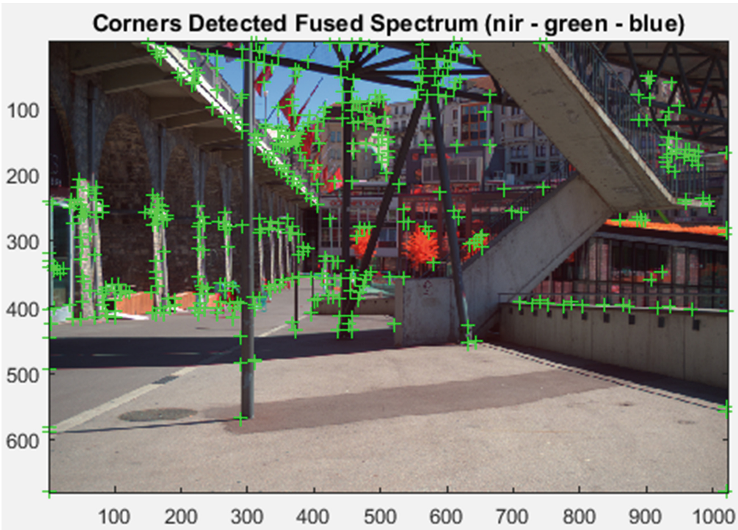

Fig. 6. Cross-Spectral image corner detection

of the feature detection is mainly due to the complex corner measure with this prune approach, according with [13], and efficient discard non corners occurs, which significantly reduces the selection and evaluation effort for the presence of

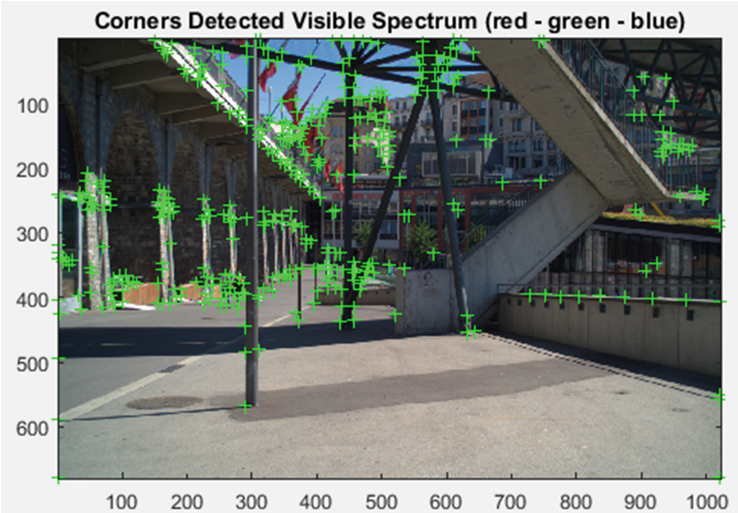

Fig. 7. RGB image corner detection

Table 1. Accuracy from the proposed approach comparing with the traditional Harris algorithm using RGB and Cross-Spectral images with all the processed image categories.

Technique	Average difference in repeatability rate			
	country	*field*	*urban*	*old-building*
Traditional Harris with RGB image	4.22	4.47	3.68	3.84
Traditional Harris with Cross-Spectral image	2.72	2.85	2.43	2.75
Pruning Harris with RGB image	2.12	2.05	2.32	2.07
Pruning Harris with Cross-Spectral image	1.68	1.75	1.76	1.31

corners to only corner like regions. In Harris, the trace (M) term is introduced so that edges can also be detected. Ignoring the $trace(M)$ term, the $det(M)$ term alone is sufficient to select corner regions. Based on their observations, they propose a pruning technique that approximates the $det(M)$, (i.e. determinant of the auto correlation matrix M) for selecting corner candidate pixel and after that they apply the conventional Harris corner measure on only the corner candidate set to extract the final corners. With this, there is not a global compute for the corner response on the entire image, hence potentially resulting in higher computation time savings.

Table 2. Average % Speed Up of from the proposed approach comparing with the traditional Harris algorithm using RGB and Cross-Spectral images with all the processed image categories.

Technique	% Speed-Up			
	country	*field*	*urban*	*old-building*
Traditional Harris with Cross-Spectral vs RGB image	28.37	29.51	28.75	23.41
Pruning Harris vs Traditional Harris with RGB image	35.72	36.68	36.15	34.95
Pruning Harris vs Traditional Harris with Cross-Spectral image	46.17	48.85	49.71	49.97
Pruning Harris with Cross-Spectral versus RGB image	62.33	64.65	67.31	63.39

Pixels that maximize $det(M)$ also have a high value for λ_2. For a high value of $det(M)$, the pixel must have a large value for a corner response (CR). We propose to choose only such pixels as corner candidates. Applying an appropriate threshold can discard pixels with low (CR) values.

4 Experimental Results

The proposed approach has been evaluated using Cross-Spectral images obtained from a fusion process (NIR, G, B), Fig. 5 shows an example of the fusion process, and also applying a pruning technique obtained from the equation presented above, The cross-spectral data set came from [18]. The *country*, *urban*, *old-building* and *field* categories have been considered for evaluating the performance of the proposed approach, examples of this dataset $country, field, urbanandold-buildingcategory$ are presented in Fig. 4. This dataset consists of 477 registered images categorized in 9 groups captured in RGB (visible spectrum) and NIR (Near Infrared spectrum). The *country* category contains 52 pairs of images of (1024×680 pixels), the *urban* category contains 58 pairs of images of (1024×680 pixels), The *old-building* category contains 51 pairs of images of (1024×680 pixels) while the *field* contains 51 pairs of images of (1024×680 pixels). In order to make the experiments 650 pairs of cross-spectral images from each of these categories has been generated from their corresponding (RGB-NIR) images. It should be noted that images are correctly registered, so that a pixel-to-pixel correspondence is guaranteed.

The parameters that we use in the Harris algorithms was a Gaussian window with W $= 3 \times 3$ kernel and $\sigma = 0.3$ as the baseline algorithms. For the proposed technique, the pruning algorithm is first applied to the entire image in order to select the corner candidate set. Next, the corner measure of the corresponding baseline algorithm is applied to extract the final corners. We compare the obtained results from the traditional Harris algorithm, using the Cross-spectral images, the original RGB images, an example of Harris detector using RGB image is showed in Fig. 7 and also applying the pruning technique in terms of the accuracy and timing. We use affine image transformations such changes in viewpoint, scale, rotation and illumination. For all images, we apply the thresholds on the corner response so that the final corner measures (λ_2 *or* R) could obtained a better accuracy as shows the table results. An example of Harris detector using Cross-Spectral image is showed in Fig. 6.

For this experiment of Harris corner evaluation we used a 3.2 eight core processor with 16 GB of memory with a NVIDIA GeForce GTX970 GPU. The accuracy of the final corners extracted is evaluated using the repeatability rate, which is defined as the number of points repeated between two images with respect to the total number of detected points. Despite of a large image data sets evaluated, a notable speedup of approximately a 30% over the traditional Harris detector is still observed, when the pruning technique is used.

Table 1 presents the results obtained with the four techniques used to evaluated the accuracy and the Table 2 shows the results for the speed-up for each category. It can be appreciated that the pruning technique reaches the best results in all cases.

5 Conclusion

This paper tackles the challenging problem of improve the feature detector algorithm, in this case, evaluating the Harris corner detector algorithm, using Cross-Spectral images, in combination with a pruning technique to obtained a better accuracy and reduce computational cost.

We have evaluated a low cost pruning technique to accelerate the Harris corner detectors by using an approximate corner indicator derived from the conventional corner measure. Evaluations for repeatability showed that the corner candidates selected by the proposed pruning technique include most of the corners found by the baseline detectors. The approximate measure used for pruning allows high thresholds to be applied to remove non corner regions, while retaining a significant amount of corners. This facilitates the selection of a small but near-complete set of corner candidates, which results in significant computation savings on corner response evaluation. Experimental results demonstrate that the proposed technique achieves significant speedup in all the experiments realized. The pruning technique is well suited for high performance and low cost embedded systems. Our future work will focus on improve the thresholding process to reach more accuracy.

Acknowledgment. This work has been partially supported by the ESPOL under projects PRAIM and KISHWAR.

References

1. Sappa, A.D., Vintimilla, B.X.: Cost-based closed-contour representations. J. Electron. Imaging **16**, 023009–023009 (2007)
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. Shaw, G.A., Burke, H.K.: Spectral imaging for remote sensing. Lincoln Lab. J. **14**, 3–28 (2003)
4. Ceamanos, X., Douté, S., Luo, B., Schmidt, F., Jouannic, G., Chanussot, J.: Intercomparison and validation of techniques for spectral unmixing of hyperspectral images: a planetary case study. IEEE Trans. Geosci. Remote Sens. **49**, 4341–4358 (2011)
5. Moroni, M., Dacquino, C., Cenedese, A.: Mosaicing of hyperspectral images: the application of a spectrograph imaging device. Sensors **12**, 10228–10247 (2012)
6. Bruce, D.A., Owens, G.P., Maliki, A., Ajeel, A.A., et al.: Capabilities of remote sensing hyperspectral images for the detection of lead contamination: a review (2012)
7. Pohl, C., Van Genderen, J.L.: Review article multisensor image fusion in remote sensing: concepts, methods and applications. Int. J. Remote Sens. **19**, 823–854 (1998)
8. Zagoruyko, S., Komodakis, N.: Learning to compare image patches via convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4353–4361 (2015)
9. Suárez, P.L., Sappa, A.D., Vintimilla, B.X.: Cross-spectral image patch similarity using convolutional neural network. In: 2017 IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM), pp. 1–5. IEEE (2017)
10. Aguilera, C.A., Aguilera, F.J., Sappa, A.D., Aguilera, C., Toledo, R.: Learning cross-spectral similarity measures with deep convolutional neural networks. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, p. 9. IEEE (2016)
11. Suárez, P.L., Villavicencio, M.: Canny edge detection in cross-spectral fused images. In: 2017 International Conference on Information Systems and Computer Science, ENFOQUE UTE (INCISCOS), pp. 16–30. LATINDEX (2017)
12. Harris, C., Stephens, M.: A combined corner and edge detector. In: Alvey Vision Conference, Manchester, UK, vol. 15, pp. 10–5244 (1988)
13. Wu, M., Ramakrishnan, N., Lam, S.K., Srikanthan, T.: Low-complexity pruning for accelerating corner detection. In: 2012 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1684–1687. IEEE (2012)
14. Wilkinson, G.G.: Results and implications of a study of fifteen years of satellite image classification experiments. IEEE Trans. Geosci. Remote Sens. **43**, 433–440 (2005)
15. Iordache, M.D., Bioucas-Dias, J.M., Plaza, A.: Sparse unmixing of hyperspectral data. IEEE Trans. Geosci. Remote Sens. **49**, 2014–2039 (2011)
16. Harris, J.L.: Image evaluation and restoration. JOSA **56**, 569–574 (1966)
17. Pavithra, A.S.: Agile segmentation and classification for hyper spectral image using Harris Corner Detector. Int. J. Sci. Res. **4**, 196–199 (2015)

18. Brown, M., Süsstrunk, S.: Multi-spectral SIFT for scene category recognition. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 177–184. IEEE (2011)
19. Dore, A., Beoldo, A., Regazzoni, C.S.: Multitarget tracking with a corner-based particle filter. In: 2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), pp. 1251–1258. IEEE (2009)