Proceedings of the2nd International Conference on Machine Vision and Machine Learning
Barcelona, Spain – July 13-14, 2015
Paper No. 331

Cross-spectral image registration and fusion: an evaluation study

Mildred Cruz[1], Cristhian Aguilera[2], Boris Vintimilla[1]
Ricardo Toledo[2] and Angel D. Sappa[1,2]

[1]Facultad de Ingeniería Eléctrica y Computación (FIEC)
Escuela Superior Politécnica del Litoral (ESPOL)
Campus Gustavo Galindo, Km 30.5 vía Perimetral
09-01-5863. Guayaquil-Ecuador
{mlcruz, boris.vintimilla}@espol.edu.ec

[2]Computer Vision Center
Edificio O, Campus UAB
08193 Bellaterra
Barcelona, España
{caguilera, ricard, asappa}@cvc.uab.es

Abstract- This paper presents a preliminary study on the registration and fusion of cross-spectral imaging. The objective is to evaluate the validity of widely used computer vision approaches when they are applied at different spectral bands. In particular, we are interested in merging images from the infrared (both long wave infrared: LWIR and near infrared: NIR) and visible spectrum (VS). Experimental results with different data sets are presented.

Keywords: multispectral imaging; image registration; data fusion; infrared and visible spectra;

1. Introduction

Recent advances in imaging sensors allow the combined usage of cross-spectral cameras to tackle classical computer vision problems (e.g., Barrera et al. (2012, 2013), Mouats (2015)). Among the most popular applications based on the combined use of cross-spectral cameras we can mention video surveillance, based on the combined use of visible and infrared cameras; or the 3D data generation using the well know Kinect system. In both cases the information provided by the camera sensors is independently used. On the contrary to the previous approaches, in the current work we propose a study to merge the information in a single representation that could be later on used for different applications. This way of facing the problem is generally used in the remote sensing field, where cross/multi-spectral images are used by merging them in a single representation. The proposed study addresses two problems. First the cross-spectral image registration, which is tackled by evaluating feature based (Aguilera et al. 2012) and mutual information based approaches (Barrera et al. 2010). Then the data fusion and image representation part is considered by evaluating different space representations (Villiers and Jermy 2013). Conclusion from evaluations with different platforms (NIR-VS and LWIR-VS) are presented.

2. Image Registration

The image registration problem has been widely studied in the literature (Zitová and Flusser 2003) and consists on overlaying two images of the same scene, taken at different times, possibly from different viewpoints and by different sensors. These images will be referred to as the source image I_s (given image) and target image I_T (reference image). In general, the registration process consists of two steps. In the first step, the correspondences between elements of the source and target images are found; then, in the second step, the global distance (error) between all the pairs of points is minimized. This process has been largely studied in the case of images belonging to the same spectral band. On the contrary, it still needs to be studied to tackle the case of images belonging to different spectral band, in particular the correspondence search step. In the current work two approaches are evaluated for finding correspondences. In the first case features points are detected and described using (Aguilera et al. 2012). In the second the mutual information value (Barrera et al. 2010) is used for finding correspondences.

2.1. Feature based approaches

Feature points detection and description are at the base of different computer vision problems. During last decade several approaches have been proposed, being the SIFT algorithm (Lowe 2001) one of the most widely used ones. Several contributions have been proposed trying to improve the performance of SIFT (i.e., speed, repeatability, etc.) (Bay et al. 2008, Calonder et al. 2010, Leutenegger et al. 2011). All the approaches mentioned before have been proposed for applications that involve images from the same spectral band, generally Visible Spectrum (VS) images. Recently, applications that combine feature points from images at different spectral band are being developed. These works are mainly based on the use of classical SIFT algorithm, or minor modifications to the classical approach. For instance (Yi et al. 2008) proposes a scale restriction criteria in order to reduce the number of incorrect matches of SIFT when it is adopted to tackle the multispectral case. Accurate matching results have been reported when the spectral bands of the pair of images are somehow near (NIR-VS) (Brown, M. and Süsstrunk 2011); however, further improvements are needed for tackling those cases where the spectral bands are far away from each other (LWIR-VS).

In this section the usage of and Edge Oriented Histogram (EOH) based approach, proposed by (Aguilera et al. 2012), is considered for finding correspondences. It consists of a scale-space pyramid, like the one used by SIFT, but by modifying the feature vector in such a way to incorporate spatial information from the contours of each keypoint without using gradient information. Contours are detected by Canny (Canny 1986) algorithm and keypoints are described by means of an edge orientation histogram computed from five directional filters (Manjunath et al. 2001). This allows us to generate a correlated parameter space in both the VS and LWIR images. This way of detecting and describing feature points has been robust to both cross spectral domains LWIR-VS and NIR-VS. More details of the implementation can be found in (Aguilera et al. 2012). Once correspondences are found registration parameters are obtained (rotation and translation) by minimizing mean square error (MSE).

Figure 1 shows a LWIR-VS pair with 20 matches used for the registration. As can be appreciated (Fig. 1 (right)) although the result is quite acceptable, there are some regions where the mismatching can be easily appreciated (see highlighted region). The main drawback of this approach lies on both the correct tuning of the algorithm (size of the windows to be described) and the amount of correct matches. Regarding the size, an 80x80 window is used in the current work (as suggested by Aguilera et al. 2012). Regarding the detection of correct matches, a RANSAC approach is considered to filter outlier. The low number of correct matches is the main limitation of feature based approaches in the multispectral domain. In the next section a mutual information based approach is presented to overcome this limitation.

Fig. 1. Registration of LWIR and VS images. (left) LWIR image with the set of keypoints (20 points); (middle) VS image with the corresponding keypoints (20 points); (right) Registration result obtained by minimizing MSE (*this figure is best viewed in color*).

On the contrary to the previous case better results are obtained when the image pair consists of a visible and a near infrared images (NIR-VS). In this cross-spectral case a larger amount of correct matches are found, which at the end results in a better registration. Figure 2 presents an illustration of the results obtained with the *oldbuilding* data set from (Brown, M. and Süsstrunk 2011). In Fig. 2 (right), it can be appreciated that after the registration the objects contained in the scene correctly overlap.

Fig. 2. Registration of NIR and VS images. (left) NIR image with selected keypoints; (middle) VS image with selected keypoints; (right) Registration result obtained by minimizing MSE over selected keypoints (*this figure is best viewed in color*).

2.2. Mutual information based approaches

In order to overcome the limitation on the number of points mentioned above the rigid registration approach presented in (Rahunathan et al. 2005) is used. This approach maximizes the Mutual Information (MI) (Barrera et al. 2010) using the Mean Square Error (MSE) between the images. The maximization of Mutual Information involves finding a transformation from the coordinate frame of the given image (I_s) to that of the target image (I_T). Figure 3 (left) and (middle) show LWIR and VS images of the same scene; the MI based registration result is presented in Fig. 3 (right). Like in the previous section, input images are overlapped using different colours to qualitatively appreciate registration results. Unfortunately, there is not a ground truth to quantitatively evaluate the accuracy of the results, but in general the registration obtained with MI is considerably better in comparison with the one obtained with the feature based approach.

The registration of NIR and VS pair of images has been also evaluated using MI; as expected, the obtained results were visually correct, in most of the cases the results were slightly better to the ones obtained using the feature based approach.

Fig. 3. (left) Input LWIR image. (middle) Input VS image. (right) Result from the MI Registration (input images are overlapped using different colours to appreciate the results).

3. Image Fusion

In this section we just explore classical fusion approaches used in the remote sensing domain when they are considered in the set of cross-spectral images evaluated in the current approach. Main difference with remote sensing images lies in the kind of objects contained in the scene (i.e., material) as well as the kind of infrared images (i.e., LWIR and NIR). In the case of NIR-VS image fusion, Normalized Difference Vegetation Index (NDVI) or Ratio Vegetation Index (RVI) can be used to easily detect vegetation in the scene. However, in the case of LWIR-VS, new *ad hoc* approaches need to be developed depending on the application. Figure 4 shows two illustrations of image fusion using the NDVI and RVI indices respectively; in these cases the LWIR image has been registered and merged with the corresponding VS. The NDVI and RVI, which have been proposed to be used with NIR images, have been computed to explore their use in other spectral band. As can be appreciated, in this particular cross-spectral scenario (LWIR-VS) these indices do not help so much to detect vegetation in the scene (glass and part of the road are also represented with the same values than vegetation). In summary, other indices need to be envisaged.

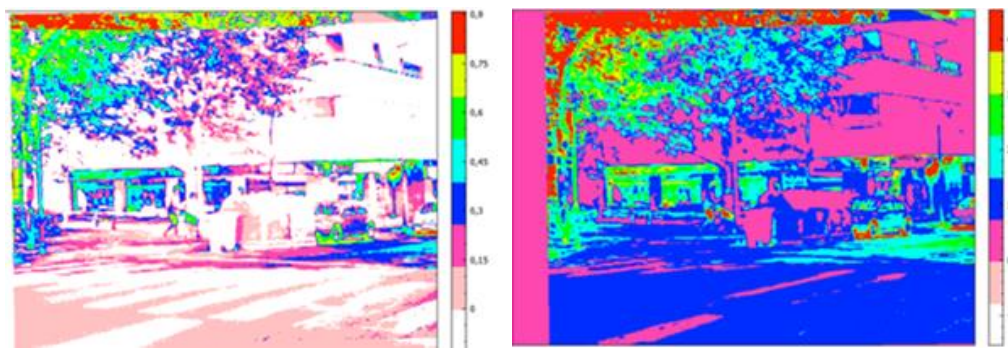

Fig. 4. (left) NDVI, using LWIR instead of NIR. (right) RVI, using LWIR instead of NIR (*this figure is best viewed in color*).

4. Conclusion

We can conclude, as in many computer vision problems, that there is not a unique pipeline valid for the different domains. Registration and fusion of cross-spectral images, that are near in the spectral domain (e.g., NIR-VS), in general can be tackled with feature based approaches. On the contrary, if the images to be registered correspond to spectral band far away from each other, it is better to use mutual information as a similarity measure. Regarding the fusion results, it depends on both the application and the kind of images. Applications in the agricultural domain, using a pair of images from the NIR and VS can benefit from the large literature in the remote sensing field (e.g., NDVI, RVI). On the contrary, if images to be merged belong to far away from each other spectral band a specific approach, depending on the application, need to be developed. As a future work, a more rigorous framework need to be envisaged to perform a quantitative evaluation that support the results obtained in the current work.

Acknowledgements

This work has been partially supported by the Spanish Government under Project TIN2014-56919-C3-2-R and PROMETEO Project of the “Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación de la República del Ecuador”. Cristhian A. Aguilera was supported by a grant from Universitat Autònoma de Barcelona.

References

Aguilera C., Barrera F., Lumbreras F., Sappa A., and Toledo R., (2012). Multispectral Image Feature Points, Sensors, Vol 12, No. 9, pp. 12661-12672.

Barrera F., Lumbreras F. and Sappa A. (2010). Multimodal Template Matching Based on Gradient and Mutual Information Using Scale Space, IEEE Int. Conf. on Image Processing, Hong Kong, pp. 2749-2752.

Barrera F., Lumbreras F. and Sappa A. (2012). Multimodal Stereo Vision System: 3D Data Extraction and Algorithm Evaluation, IEEE Journal of Selected Topics in Signal Processing, Vol. 6, No. 5, pp. 437-446.

Barrera F., Lumbreras F. and Sappa A. (2013).Multispectral Piecewise Planar Stereo using Manhattan-World Assumption, Pattern Recognition Letters, Vol. 34, No.1, pp. 52-61.

Bay, H., Tuytelaars, T. and Van Gool, L. (2008). Speeded Up Robust Features, Computer Vision and Image Understanding, Vol. 110, No. 3, pp. 346-359.

Brown, M. and Süsstrunk, S. (2011). Multi-spectral SIFT for scene category recognition, Int. Conference on Computer Vision and Pattern Recognition, Colorado Springs, USA, pp. 177-184.

Calonder, M., Lepetit, V., Strecha, C. and Fua P. (2010). BRIEF: binary robust independent elementary features, European conference on Computer vision, 2010, pp. 778-792.

Canny, J. A. (2986). Computational approach to edge detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, pp. 679–698.

Leutenegger, S., Chli, M. and Siegwart, R. (2011). BRISK: Binary Robust invariant scalable keypoints, Int. Conference on Computer Vision, Washington, DC, USA, pp. 2548-2555.

Lowe, D.G. (2001). Local Feature View Clustering for 3D Object Recognition, Int. Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, pp. 682–688.

Manjunath, B., Ohm, J., Vasudevan, V., Yamada, A., (2001), Color and texture descriptors, IEEE Trans. Circuits Systems and Video Technology, Vol. 11, pp. 703–715.

Mouats T., Aouf N., Sappa A., Aguilera C. and Toledo R. (in press) Multi-Spectral Stereo Odometry, IEEE Transactions on Intelligent Transportation Systems.

Rahunathan, Smriti, D. Stredney, P. Schmalbrock, and B.D. Clymer (2005). Image Registration Using Rigid Registration and Maximization of Mutual Information. Poster presented at: MMVR13. The 13th Annual Medicine Meets Virtual Reality Conference; Long Beach, CA.

Ricaurte P., Chilán C., Aguilera-Carrasco C., Vintimilla B. and Sappa A. (2014). Feature Point Descriptors: Infrared and Visible Spectra, Sensors, Vol. 14, No. 2, pp. 3690-3701.

Villiers J. and Jermy R. (2013). An analysis of fusion algorithms for LWIR and visual images, 24th Annual symposium of the Pattern Recognition Association of South Africa.

Zitová B. and Flusser J. (2003). Image registration methods: A survey, Image and Vision Computing, Vol. 21, No. 11, pp. 977–1000.

Yi, Z., Zhiguo, C., and Yang, X. (2008). Multi-spectral remote image registration based on SIFT, Electronics Letters, Vol. 44, 107–108.