

Thesis Title	Multispectral video analysis for deformable texture recognition using a snapshot camera	
Director	Ludovic MACAIRE	E-mail: ludovic.macaire@univ-lille.fr
Co-supervisor	Olivier LOSSON	E-mail: olivier.losson@univ-lille.fr
Co-supervisor	Benjamin MATHON	E-mail: benjamin.mathon@univ-lille.fr
Co-supervisor	Clémence Prévost	clemence.prevost@univ-lille.fr
Laboratory	CRISTAL UMR CNRS 9189	https://www.cristal.univ-lille.fr
Team	Imagerie couleur	http://color.univ-lille.fr/
Funding		

Application only on https://adum.fr/as/ed/voirproposition.pl?site=adumR&matricule_prop=52649#version

Thesis subject

This thesis is about plant recognition in outdoor scenes. It extends Anis Amziane's doctoral work about weed identification by multispectral image analysis [1]. The camera technology that was used to acquire the analyzed images requires the observed plants to be motionless during acquisition [2]. This is a strong limitation for applications in outdoor since plants move due to the wind, in such a way that their leaves and stems can be seen as time-varying textural surfaces. This thesis aims at developing methods to automatically identify such deformable textures by image sequence (i.e., video) analysis.

Classical methods for motion identification by video analysis try to estimate the motion of each surface element of the observed objects. During the movement of an object or the deformation of a texture, each moving surface element in the scene is projected onto different pixels in the successive images acquired by the camera. The main objective is then to retrieve the spatial coordinates of the pixel associated to the same given surface element in each of the images of the sequence. A common assumption to achieve this goal is that the energy reflected by each surface element is constant throughout the object displacement, hence the pixel associated to this element is also constant over all the image sequence. Classical motion analysis methods therefore use correlation measures or optical flow approaches to estimate motion [3].

Plant surfaces have deformable textures that can be observed in several spectral domains. Though the visible domain is easily accessible, near infrared allows one for a more efficient distinction of different plant species, such as weeds and crops. However, classical cameras (either gray-level or color ones) cannot acquire image sequences in the infrared domain.

Several snapshot multispectral cameras have been proposed in the literature in past years. Such scanless cameras acquire multispectral images at video rate. The so-called

(M)SFA (*multi-spectral filter array*) technology has recently attracted the interest of the image processing community [4,5]. It is a generalization of the ground concept of single-sensor color cameras, namely a compromise between spatial and spectral resolutions. The sensor is covered by a mosaic of spectral filters, each of which being associated to a pixel and sensitive to a narrow spectral band. A single spectral component is therefore available at each pixel of the resulting image, such that this so-called raw (or MSFA) image has the same storage cost as a gray-level image. A multispectral video delivered by such a snapshot camera is thus made of a sequence of raw images acquired at video rate and stored as a gray-level image set.

The spectral component values that are missing in a raw image can then be estimated by demosaicing to reconstruct a fully-defined multispectral image [6,7]. The quality of the estimated multispectral image depends on the SFA, on the number of spectral bands, on their nature, and on the demosaicing procedure. Because demosaicing causes errors on the estimated values of missing components, this process degrades the quality of the texture representation with regards to raw images. Furthermore, a sequence of fully-defined multispectral images is highly memory greedy. We therefore propose to directly analyze raw image sequences to identify deformable textures. To our knowledge, very few works deal with multispectral video analysis [4,8,9]. This doctoral subject then tackles original open problems.

Motion identification by raw image analysis cannot rely on the hypothesis of gray-level constancy of the pixel associated to a given surface element in the successively acquired images. Indeed, a same moving surface element is likely to be presented by neighboring pixels in two successively acquired images, hence by raw values that are associated to different spectral bands. As a result, nor multi-scale approaches nor phase-based ones in the spatial frequency domain can be directly used to extract motion from raw image sequences.

Such raw image sequences will have to be analyzed by taking the SFA pattern of the camera into account. A first way is to proceed sparsely and to exploit raw images for keypoint detection [10], then to match the detected keypoints on the successive raw images. A second approach is to follow a dense strategy and to take inspiration from texture descriptors adapted to the SFA pattern to analyze textures from raw images, as designed by the Imagerie Couleur team [11]. Similarly, the challenge will be to design descriptors that are adapted to deformable textures that change over time.

Thesis planning

M1-M6 : Bibliography about motion estimation by classical video analysis methods and about formation and analysis of multispectral raw image sequences.

M7-M12 : Motion estimation by local correlation measures within raw multispectral image sequences.

M13-18 : Motion estimation by optical flow-based and phase-based analyses of multispectral videos.

M19-M24 : Design of texture descriptors by mono- and multi-scale analyses of raw image sequences. Tests on raw images either simulated from fully-defined multispectral images

or acquired by IMEC snapshot mosaic camera available at the EquipEx IrDIVE-Continuum.
M25-M30 : Application to the recognition of moving plants.
M31-M36 : Manuscript writing.

References

- [1] Anis Amziane. Texture features extracted from multispectral images acquired under uncontrolled illumination conditions—Application to precision farming. PhD thesis, Oct. 2022, <https://www.theses.fr/s323059>.
- [2] Anis Amziane, Olivier Losson, Benjamin Mathon, Aurélien Duménil, and Ludovic Macaire. Reflectance Estimation from Multispectral Linescan Acquisitions under Varying Illumination—Application to Outdoor Weed Identification, *Sensors*, 2021, 21 (11), 3601, <https://dx.doi.org/10.3390/s21113601>
- [3] Cédric Marinel, Benjamin Mathon, Olivier Losson, Ludovic Macaire. Comparison of Phase-based Sub-Pixel Motion Estimation Methods , *Procs. IEEE International Conference on Image Processing (ICIP)*, Oct. 2022, Bordeaux, France. pp.561-565, <https://doi.org/10.1109/ICIP46576.2022.9897338>
- [4] Kathleen Vunckx and Wouter Charle. Accurate Video-Rate Multi-Spectral Imaging Using IMEC Snapshot Sensors, *Procs. 11th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2021, pp. 1-7, <https://doi.org/10.1109/WHISPERS52202.2021.9483975>
- [5] Ying Shen, Jie Li, Wenfu Lin, Liqiong Chen, Feng Huang, and Shu Wang. Camouflaged Target Detection Based on Snapshot Multispectral Imaging. *Remote Sensing*, 2021, 13 (19), 3949, <https://doi.org/10.3390/rs13193949>
- [6] Sofiane Mihoubi, Olivier Losson, Benjamin Mathon, and Ludovic Macaire. Multispectral demosaicing using pseudo-panchromatic image, *IEEE Transactions on Computational Imaging*, 2017, 3 (4), pp. 982-995, <https://doi.org/10.1109/TCI.2017.2691553>
- [7] Grigorios Tsagkatakis, Maarten Bloemen, Bert Geelen, Murali Jayapala, and Panagiotis Tsakalides, Graph and Rank Regularized Matrix Recovery for Snapshot Spectral Image Demosaicing, *IEEE Transactions on Computational Imaging*, 2019, 5 (2), pp. 2333-9403, <https://doi-org/10.1109/TCI.2018.2888989>
- [8] Raffaele Vitale, Cyril Ruckebusch, Ingunn Burud, and Harald Martens. Hyperspectral Video Analysis by Motion and Intensity Preprocessing and Subspace Autoencoding, *Frontiers in Chemistry*, 2022, <https://doi.org/10.3389/fchem.2022.818974>
- [9] Lulu Chen, Yongqiang Zhao, and Seong G. Kong. SFA-guided mosaic transformer for tracking small objects in snapshot spectral imaging, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2023, pp. 223–236, <https://doi.org/10.1016/j.isprsjprs.2023.09.015>
- [10] Xiangyu Zhang, Ling Zhang, and Xin Lou. A Raw Image-Based End-to-End Object Detection Accelerator Using HOG Features, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2022, 69 (1), pp. 322-333, <https://doi.org/10.1109/TCSI.2021.3098053>
- [11] Sofiane Mihoubi, Olivier Losson, Benjamin Mathon, and Ludovic Macaire. Spatio-spectral Binary Patterns Based on Multispectral Filter Arrays for Texture Classification. *Journal of the Optical Society of America A*, 2018, 35 (9), pp.1532-1542,

<https://doi.org/10.1364/JOSAA.35.001532>

Titre Thèse	Analyse de vidéos multi-spectrales pour la reconnaissance de textures déformables observées par une caméra de type snapshot	
Directeur	Ludovic MACAIRE	E-mail : ludovic.macaire@univ-lille.fr
Co-encadrant	Olivier LOSSON	E-mail :olivier.losson@univ-lille.fr
Co-encadrant	Benjamin MATHON	E-mail :benjamin.mathon@univ-lille.fr
Co-encadrante	Clémence Prévost	clemence.prevast@univ-lille.fr
Laboratoire	CRISTAL UMR CNRS 9189	https://www.cristal.univ-lille.fr
Equipe	Imagerie couleur	http://color.univ-lille.fr/
Financement prévu		

Candidature seulement sur https://adum.fr/as/ed/voirproposition.pl?site=adumR&matricule_prop=52649#version

Résumé du sujet de thèse

Cette thèse concerne la reconnaissance de plantes en scènes extérieures. Elle fait suite à la thèse d'Anis Amziane qui a porté sur l'identification d'adventices par analyse d'images multi-spectrales [1]. La technologie de la caméra utilisée dans cette thèse impose que les plantes observées doivent être fixes lors de l'acquisition des images [2]. Cela limite la portée applicative de ces travaux en milieu extérieur car les plantes se déforment alors sous l'effet du vent, leurs feuilles et tiges constituant des surfaces texturées qui se déforment au cours du temps. Cette thèse vise à développer des méthodes permettant d'identifier automatiquement ces textures déformables par l'analyse de séquences temporelles d'images, c'est-à-dire de vidéos.

Les méthodes classiques de détermination du mouvement par analyse de vidéos tentent de déterminer le déplacement de tous les éléments constituant les surfaces des objets. Lors du mouvement de l'objet ou de la déformation de la texture, chaque élément de surface mobile dans la scène va se projeter sur des pixels différents dans les images acquises consécutivement par la caméra. L'objectif principal est donc de retrouver les coordonnées spatiales du pixel associé à un même élément de surface dans chacune des images composant la séquence. Pour ce faire, on suppose que l'énergie réfléchie par chaque élément de surface se conserve au cours du mouvement de l'objet. Cela se traduit classiquement par la constance du niveau de gris du pixel associé à cet élément de surface dans chacune des images de la séquence. Les méthodes classiques d'analyse de mouvement procèdent alors par des mesures de corrélation, ou développent des approches de flot optique pour déterminer le mouvement [3].

Les surfaces des plantes présentent des textures déformables qui peuvent être observées dans différents domaines spectraux. Si le domaine visible est classiquement accessible, le proche infra-rouge présente l'avantage de permettre de séparer plus efficacement

différentes espèces végétales, donc adventices et cultures. Or les caméras classiques (en niveaux de gris ou couleur) ne permettent pas d'acquérir des vidéos ou séquences d'images dans le domaine du proche infra-rouge.

Différentes caméras multispectrales de type snapshot (instantané) ont été proposées dans la littérature ces dernières années. Ces caméras sans balayage acquièrent des images multispectrales à cadence vidéo. La technologie dite (M)SFA (*(multi-)spectral filter array*) a récemment attiré l'intérêt de la communauté scientifique [4,5]. Il s'agit d'une généralisation du concept de caméra couleur matricielle à l'imagerie multispectrale, à savoir un compromis entre les résolutions spatiale et spectrale. Le capteur est pour cela recouvert d'une mosaïque de filtres spectraux, dont chacun est associé à un pixel et n'est sensible qu'à une bande spectrale réduite. Dans l'image résultante (appelée image brute (raw) ou (M)SFA), une seule composante spectrale est donc disponible pour un pixel donné. Ainsi, la place mémoire occupée par l'image brute est identique à celle d'une image en niveaux de gris. Une vidéo multi-spectrale délivrée par ce type de caméra est donc constituée d'une séquence d'images brutes acquises à la cadence vidéo classique et stockée comme une séquence d'images brutes codées en niveaux de gris.

Les composantes spectrales manquantes au sein de chaque image brute peuvent ensuite être estimées par dématricage [6,7] afin d'obtenir une vidéo multispectrale pleinement définie. La qualité d'une telle vidéo est dépendante du SFA, du nombre de bandes spectrales, de leur nature et de la méthode de dématricage appliquée à chaque image. Comme le dématricage induit des erreurs sur les valeurs estimées des composantes manquantes, il dégrade la qualité de représentation des textures par rapport aux images brutes. Par ailleurs, une séquence d'images multi-spectrales pleinement définies est particulièrement gourmande en taille mémoire. Par conséquent, nous proposons d'analyser directement les séquences d'images brutes pour reconnaître les textures déformables. À notre connaissance, très peu de travaux portent actuellement sur l'analyse de vidéos multi-spectrales [4,8,9]. Les problèmes abordés par ce sujet de thèse sont donc tout à fait originaux.

Les techniques de détermination du mouvement par analyse d'images brutes ne peuvent pas s'appuyer sur les hypothèses de constance du niveau de gris représentant le pixel associé à un élément de surface dans chacune des images acquises consécutivement. En effet, comme la seule composante spectrale disponible est différente entre pixels voisins spatialement, leurs niveaux sont faiblement corrélés quand ils sont associés à des éléments de surface différents. Par conséquent, les niveaux des pixels qui observent, dans les images acquises consécutivement, un même élément de surface en mouvement, dépendent des composantes spectrales disponibles en ces pixels, et sont donc différents entre eux. Les approches multi-échelles spatiales ne peuvent donc pas être appliquées directement puisque les composantes spectrales disponibles dépendent du SFA. Les approches procédant par analyse de la phase dans le domaine des fréquences spatiales ne peuvent pas non plus être appliquées directement pour extraire le mouvement à partir

de séquences d'images brutes.

Les méthodes d'analyse de ces séquences d'images brutes devront inévitablement tenir compte du motif du SFA équipant la caméra. Une première piste de travail consiste à exploiter les images brutes pour la détection de points clés [10]. Cette approche éparse consistera à mettre en correspondances les points clés détectés sur les images brutes successives. Une seconde piste suivant une stratégie dense est de s'inspirer des descripteurs de texture conçus par les membres de l'équipe Imagerie Couleur pour l'analyse de textures fixes et adaptés au treillis SFA [11]. Il s'agira, sur un principe similaire, d'imaginer des descripteurs dédiés aux textures déformables dans le temps.

Planning de la thèse

M1-M6 : Bibliographie sur l'analyse du mouvement par méthodes classiques d'analyse de vidéos et sur la formation et l'analyse de séquences d'images brutes multi-spectrales.

M7-M12 : Détermination du mouvement par mesure de corrélations locales au sein des séquences d'images brutes multi-spectrales.

M13-18 : Étude du mouvement par flot optique s'appuyant sur la phase au sein des vidéos multi-spectrales.

M19-M24 : Développement de descripteurs de texture par analyse mono-échelle et multi-échelle de séquence d'images brutes. Expériences sur séquences d'images brutes simulées dans un premier temps à partir d'images pleinement définies. Elles pourront également être acquises par les caméras saphost mosaic de la société IMEC disponibles au sein de l'EquipEx IrDIVE-Continuum.

M25-M30 : Application à la reconnaissance de plantes en mouvement.

M31-M36 : Rédaction du manuscrit.

Bibliographie

[1] Anis Amziane. Attributs de texture extraits d'images multispectrales acquises en conditions d'éclairage non contrôlées — Application à l'agriculture de précision. Thèse soutenue en octobre 2022, <https://www.theses.fr/s323059>.

[2] Anis Amziane, Olivier Losson, Benjamin Mathon, Aurélien Duménil, and Ludovic Macaire. Reflectance Estimation from Multispectral Linescan Acquisitions under Varying Illumination— Application to Outdoor Weed Identification, *Sensors*, 2021, 21 (11), 3601, <https://dx.doi.org/10.3390/s21113601>

[3] Cédric Marinel, Benjamin Mathon, Olivier Losson, Ludovic Macaire. Comparison of Phase-based Sub-Pixel Motion Estimation Methods , *Procs. IEEE International Conference on Image Processing (ICIP)*, Oct. 2022, Bordeaux, France. pp.561-565, <https://doi.org/10.1109/ICIP46576.2022.9897338>

[4] Kathleen Vunckx and Wouter Charle. Accurate Video-Rate Multi-Spectral Imaging Using IMEC Snapshot Sensors, *Procs. 11th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2021, pp. 1-7, <https://doi.org/10.1109/WHISPERS52202.2021.9483975>

[5] Ying Shen, Jie Li, Wenfu Lin, Liqiong Chen, Feng Huang, and Shu Wang. Camouflaged Target Detection Based on Snapshot Multispectral Imaging. *Remote Sensing*, 2021, 13 (19), 3949, <https://doi.org/10.3390/rs13193949>

- [6] Sofiane Mihoubi, Olivier Losson, Benjamin Mathon, and Ludovic Macaire. Multispectral demosaicing using pseudo-panchromatic image, *IEEE Transactions on Computational Imaging*, 2017, 3 (4), pp. 982-995, <https://doi.org/10.1109/TCI.2017.2691553>
- [7] Grigorios Tsagakatakis, Maarten Bloemen, Bert Geelen, Murali Jayapala, and Panagiotis Tsakalides, Graph and Rank Regularized Matrix Recovery for Snapshot Spectral Image Demosaicing, *IEEE Transactions on Computational Imaging*, 2019, 5 (2), pp. 2333-9403, <https://doi-org/10.1109/TCI.2018.2888989>
- [8] Raffaele Vitale, Cyril Ruckebusch, Ingunn Burud, and Harald Martens. Hyperspectral Video Analysis by Motion and Intensity Preprocessing and Subspace Autoencoding, *Frontiers in Chemistry*, 2022, <https://doi.org/10.3389/fchem.2022.818974>
- [9] Lulu Chen, Yongqiang Zhao, and Seong G. Kong. SFA-guided mosaic transformer for tracking small objects in snapshot spectral imaging, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2023, pp. 223–236, <https://doi.org/10.1016/j.isprsjprs.2023.09.015>
- [10] Xiangyu Zhang, Ling Zhang, and Xin Lou. A Raw Image-Based End-to-End Object Detection Accelerator Using HOG Features, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2022, 69 (1), pp. 322-333, <https://doi.org/10.1109/TCSI.2021.3098053>
- [11] Sofiane Mihoubi, Olivier Losson, Benjamin Mathon, and Ludovic Macaire. Spatio-spectral Binary Patterns Based on Multispectral Filter Arrays for Texture Classification. *Journal of the Optical Society of America A*, 2018, 35 (9), pp.1532-1542, <https://doi.org/10.1364/JOSAA.35.001532>