Computers and Electronics in Agriculture 227 (2024) 109509

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Cost-efficient algorithm for autonomous cultivators: Implementing template matching with field digital twins for precision agriculture

Luca De Bortoli [a], Stefano Marsi [a,*], Francesco Marinello [b], Paolo Gallina [c]

[a] Image Processing Laboratory (IPL), University of Trieste, Engineering and Architecture department, via A. Valerio 10, Trieste, Italy
[b] Biosystems engineering , University of Padova, Department of Land, Environment, Agriculture and Forestry, viale Dell'Universita' 16, Legnaro (PD), Italy
[c] Applied Mechanics for Machinery, University of Trieste, Engineering and Architecture department, via A. Valerio 10, Trieste, Italy

ARTICLE INFO

Keywords:
Precision agriculture
Cultivator
Crop row detection
Machine vision
Real-time tracking
Digital twin

ABSTRACT

The paper focuses on the development of a vision system to automate the position control of a cultivator used for crop weeding. The vision algorithm allows monitoring of the cultivator's misalignment with respect to crop rows, with real-time processing. The key content includes the introduction of a self-generated digital twin of the field model for numerical validation of different computer vision solutions and a comparison of three vision algorithms for measuring deviation. The objectives of the study are to improve the precision of misalignment measurements and ensure safe and accurate movement of the cultivator. The rationale behind the study is to address constraints such as camera installation and crop color, and to emphasize the importance of a confidence estimation feature for accurate measurement. The paper also provides an overview of related works in the literature, highlighting the two phases of plant identification and deviation measurement. Tests carried out on soybean and maize crops demonstrate the improvements allowed by the proposed algorithm in terms of higher measurement precision, even in the presence of high weed infestation or a significant number of missing plants. Additionally, the paper suggests analysis simplifications to enhance the algorithm's speed while maintaining satisfactory measurement accuracy.

1. Introduction

The weeding operation of crops is essential for soil aeration, water drainage, and removing unwanted herbs. Traditionally, trailed cultivators with furrows spaced according to the inter-row distance are implemented. However, aligning the cultivator with the sown rows require manual correction by the tractor operator or by a second operator on the cultivator.

This paper aims to study a vision algorithm able to monitor the displacement of a cultivator from the crop rows. The output of the measurement is a displacement/correction signal that is likely implementable to let an on the go electromechanical or pneumatic control of cultivator movement, allowing a single operator to concentrate exclusively on tractor steering within the sown field. Some alternative solutions have been recently published to this end by different research groups. In some case the approach to trajectory definition relies on GNSS mapping of the tractor's movement during the seeding phase (see e.g. Machleb et al. (2020)). Other researchers opt for the identification of single plants in the row through specifically trained convolutional neural networks (Liu et al., 2024; Zheng et al., 2023) or implementing other specific segmentation techniques (Wang et al., 2022). Differently from previous bibliography, the present paper introduces a novel computer vision approach based on pattern matching, which provides real-time measurement of the cultivator's misalignment with respect to the crop rows. The paper focuses on the development, testing and optimization of the misalignment measurement system, to be implemented as a feedback for the mechanical displacement actuator of the cultivator. Testing of mechanized operation is not included in the analysis; a system is hypothesized with a self-propelled chassis only in a transverse direction with respect to the tractor's travel, with numerical movement, capable of measuring the position of the mobile chassis, with a closed loop digital control that takes the deviation measurements produced by the proposed algorithms as an error function.

The system assumes the use of a front-mounted monocular camera, fixed to the cultivator's frame. The camera captures the field before cultivation, providing a coarse central perspective as the tractor usually runs parallel to the crop rows. By analyzing the video stream, the system determines the position of the crop rows relative to the fixed frame, enabling the calculation of the cultivator's deviation from the correct sowing position, typically within a few tens of centimeters. This information is crucial for the final mechanical positioning system.

* Corresponding author.
E-mail address: marsi@units.it (S. Marsi).

https://doi.org/10.1016/j.compag.2024.109509
Received 10 October 2023; Received in revised form 6 September 2024; Accepted 26 September 2024
Available online 8 October 2024

Several constraints are assumed for the measurement system's development, including knowledge of the coarsely defined inter-row distance, fixed camera installation geometry, parallel movement of the camera with the tractor, and distinguishable color of the crop from the soil. Additionally, the proposed system is designed to provide a confidence estimation, allowing the control system to weigh the misalignment measure differently, particularly in critical areas.

The proposed study is focused on maize and soybean plants (Egli, 2023). These crops were considered to be properly representative in terms of cultivated area (they are respectively the most widespread cereal and oilseed crops), inter-row spacing (normally 0.75 and 0.45 m), plant density (7–10 plants/m^2 and 40–50 plants/m^2 respectively) and also size and shape of the plant during the first cultivation phases. Additionally spacing and density pose relevant issues in the management of weeds (Singh et al., 2020).

Real-world scenarios may include areas with crossing lines from the sowing phase or sections where the sowing pattern is unrecognizable. In these situations, the confidence index is expected to provide a low value, allowing the control system to temporarily deactivate the movement system.

The following sections present similar works found in the literature, highlighting how the vision problem can be divided into two phases: plant identification and measurement of deviation from the sowing reference.

1.1. Color segmentation - Related works

Having assumed the use of a single monocular camera, neither a direct mapping of the depths nor an estimate of it is possible since, by hypothesis, the displacement occurs longitudinally to the framed plane. Having hypothesized that the plants have a known color different from the soil, their identification therefore passes through a color segmentation.

Typically the background is brown–gray and in most cases the plants to be recognized are green, a color similar however to weeds. Some crops, on the other hand, have a red–purple color.

An intuitive approach is to use a color space different from the acquisition RGB, a space on which a rule can be easily defined to filter the image. For example, it is possible to use the HSV space, as proposed in (Li et al., 2009), or the Lab space, as proposed in Correa et al. (2011) and Aden Darge (2019), to select only the pixels with the desired hue: for both proposed color spaces, the chromatic variation due to the different lighting of the scene it affects a different channel from those used to segment according to color, making these solutions more robust to lights and shadows.

As already discussed in several works, some color conversions are of interest for this purpose to convert an RGB image into gray levels in which the subjects to be identified are emphasized. The useful methods used to recognize the green on the soil are: '2G-R-B' (Søgaard and Olsen, 2003), an integer combination pixel by pixel of the values of the three RGB channels, 'Excess Green' (ExG = 2g-r-b) (Woebbecke et al., 1994), where the color coordinates r,g,b are identified for each pixel by normalization (so r+g+b = 1), and 'Modify ExG' (MExG = 1.262g-0.884r-0.311b) (Burgos-Artizzu et al., 2011), where the values of the chromatic coordinates are weighted according to coefficients identified by a genetic algorithm optimization on a dataset. Also in this case the ExG and MExG methods, which normalize the brightness pixel by pixel, are more robust to the different lighting conditions of the scene. Another identified approach consists in using a near infrared bandpass filter after camera lens (Tillett et al., 2002) thus emphasizing the contrast between the plant and the soil, regardless of their colors, using a simply monochrome camera. Similarly, multispectral input systems can be evaluated.

The segmentation of the image in the HSV and Lab color spaces can be obtained simply by applying one or more thresholds to the color values, which can be defined a prior if the lighting conditions are controlled and know. The other proposed approaches, on the other hand, produce a grayscale image where the texture to be recognized is emphasized but need to identify optimal threshold values for their segmentation. In Zong et al. (2020) propose a study on maize plants in which they compare six most common segmentation algorithms, identifying the most accurate and fastest the minimum cross entropy (Li and Lee, 1993) and Otsu (Otsu, 1979) algorithms. Among the works analyzed, the Otsu method is the most used for identifying the optimal segmentation threshold since it is assumed that to distinguish two well separated color classes (the soil and the vegetation).

In Burgos-Artizzu et al. (2011) propose a computationally simple method for calculating the threshold with respect to the Otsu algorithm, that is, they calculate the average intensity of an image converted with the MExG method to threshold it. In Montalvo et al. (2012) instead propose a double segmentation, each applying the Otsu method: with the first threshold of the ExG image, the soil is separated from all that is green; with this result the ExG image is masked and re-segmented in order to separate the shades of green to distinguish the crop from the weeds.

The work (Li et al., 2020) compares different segmentation algorithms evaluating them in terms of recognition accuracy on different crops and different lighting conditions. The results show that the segmentation method based on the Lab color space allows obtaining, on average, good results on all the cases analyzed.

Finally, image processing techniques such as filtering or morphological operators are often proposed to refine the boundary obtained from the segmentation. Further proposals to improve the recognition of rows with respect to weeds can be found in Burgos-Artizzu et al. (2011) where they propose a temporal logic product of the last ten segmented frames.

1.2. Seeding pattern estimation — related works

Having obtained a binary image represented the crop rows, the search for the seeding pattern to estimate the misalignment of the cultivator was addressed in different methods with 2D camera. Many identified works aim to look for the dominated lines of the binary image, that is, the lines that define each crop row.

The popular tool for this purpose is the Hough transform (Fisher et al., 2003): before its calculation, the edge of the binary image must be extracted. In Correa et al. (2011) identifying the edge detection algorithm that maximizes the accuracy of the lines identified with the Hough transform, suggesting to use the Roberts (Roberts, 0000) algorithm. The work (Nicholls and Green, 2017) instead proposes a method of filtering the different lines identified with the transformation, considering only those with an almost vertical direction and joining similar lines. To reduce the computational load of the transform (Ji and Qi, 2011) propose the use of its variant, the Random Hough transform, observing a temporal speed up and more accuracy with sparse plants. In Winterhalter et al. (2018), instead, propose a variant of the transform, the Pattern Hough transform, based on the recognition of the set of equidistant parallel lines: note the parallel spacing to be searched for, the proposed algorithm allows to identify angle and lateral offset of the identified plot. A last variant identified is in the work of Ericson and Åstrand (2010) where they use an omnidirectional camera (camera with fisheye or catadioptric lens) and identify the rows in the space of the Hough transform calculated in spherical coordinates.

Another approach for the search of the dominant lines passes from the analysis of the vertical histograms of the binary image. In García-Santillán et al. (2018), Jiang et al. (2015) and Romeo et al. (2012) the binary image is divided into horizontal bands on which are computed the vertical histograms and finally are analyzed for identify the position of the peaks: the set of points identified represent the center of gravity of the rows in each horizontal band. From these centers of gravity (García-Santillán et al., 2018; Jiang et al., 2015) look for the equation of the lines representing the lines by regression while Romeo

et al. (2012) calculates for each lower barycenter all the possible lines connected with the upper barycentres to select only the one that mostly cover the green area.

Okamoto et al. (2001) instead proposes a preliminary perspective correction of the acquired image from which, once binarized according to color, it is split into horizontal bands, find the peaks of crop rows and calculate by regression the dominant lines from which estimate rotation and offset of the captured sowing pattern.

Always analyzing the vertical histograms of the division into horizontal bands of the binary image (Vidović et al., 2016; Søgaard and Olsen, 2003) they also take into consideration the geometry of the system, calculating in advance the expected position of the seeding pattern to better identify the deviation. Søgaard and Olsen (2003) breaks the vertical histograms according to the sowing period and sum the various pieces of histograms in order to obtain the average histogram of the period: on this signal it searches for the position of the maximum to identify the deviation of the rows with respect to the calculation reference. Vidović et al. (2016) instead creates periodic signals at three levels, for each horizontal band, representing the ideal trend of the crop rows according to the perspective in which they are taken: each signal is correlated with the corresponding vertical histogram in order to identify the best horizontal shift.

A last variant identified for search the seeding pattern deviation is in the works (Hague and Tillett, 2001; Tillett et al., 2002) where they propose the use of a bandpass filter derived from the period of the seeding geometry: they divide the image into horizontal bands, filter the histogram of each band by setting the center band corresponding to the seeding period in perspective and finally find the displacement by look for the maximum of the filtered signal.

Eventually the digital reconstruction of plant rows, or utilizing a few representative rows, has become prevalent in recent studies, including those published in Computers and Electronics in Agriculture. This method is favored for its flexibility and ability to simulate different field conditions and complexities, especially during training phases Ban et al. (2024), Zhang et al. (2022) and de Silva et al. (2024).

2. Materials and methods

As explained in the introduction, together with the measurement of the deviation of the seeding pattern we want a confidence index capable of discriminating the capture of anomalous field. In the identified works this information is often overlooked. For example, with the proposals based on the identification of dominant lines, a confidence index could be evaluate by analyzing how many lines were detected along certain directions. With the methods that calculate vertical histograms, instead, it is possible define a confidence index by analyzing the histograms: they are a periodic functions with maxima and minima defined by the presence or absence of vegetation respectively; both from the analysis of the period of this function and from the variance between maximums and minimums it is possible to numerically estimate the framing of an ideal or less sowing pattern. In particular, the histogram already represents a vertical average of the framed vegetation therefore weeds and missing plants, if with limited extension, bring a low interference to the signal representative of the crop rows: it is possible to hypothesize a maximum modulation of the histogram for ideal sowing patterns to go down to minimum or zero values of modulations in the case of capture of strongly degraded or absent patterns.

Furthermore, in the reported works, except (Ericson and Åstrand, 2010) that it used a heavily distorted lens, the possibility of operating on distorted images due to the use of wide-angle lenses is not taken into consideration; works that analyze the image in horizontal bands they discard the most distorted ones at the base. A wide angle lens can be useful to limit the installation height of the camera, and therefore the vibrations induced on it, either to frame more than a couple of crop rows or to capture crops with high spacing. In these cases, making a single vertical histogram of the image leads to misalignments of the data which would require a geometric correction of the captured frame, that is a time consuming operation. On the other hand, if the calculation of the vertical histograms is performed on several bands, the error is minimized even if a function with some alteration of the period is obtained: in this case, however, the computation complexity is greater than the analysis of a single histogram because the period analysis operations are performed several times.

In order to develop a system for measure the deviation of seeding pattern together with a confidence index, that works in real time, with the aim of minimizing the calculations to be repeated for each image and taking into account the hypothesized working constrains (fixed and defined camera, known sowing geometry) we proposes an alternative method to those identified in the literature based on template matching: pre-calculate the ideal sowing pattern, taking into account the volume of plants, the perspective and the distortion of the camera lens, in order to identify its displacement in the video stream. In the realistic hypothesis of shooting and running parallel to the crop rows, the search for the position of the reference pattern occurs in a single dimension, i.e. across the rows; operating with a one-dimensional correlation thus greatly limiting the computation. With this method, the measurement is expected to be feasible even if only partially capturing two rows of plants up to the limit of taking a single row. Furthermore, it is possible to optimize the calculation of the correlation index using binary operators as it operates on two binary images, the acquired frame segmented according to color and a reference frame that can be defined with only two levels. The proposed algorithm will be detailed below.

To validate the proposal, two algorithms are compared in terms of measurement precision of the deviation: a literature algorithm based on the vertical histogram has been taken as a reference, some of its variants have been proposed in order to improve the measurement accuracy. The analyzed proposals are summarized below, which we will explain in the following paragraphs.

1. Reference: algorithm that calculates deviation of the crop rows on 10 horizontal bands, analyzing 10 vertical histograms, as described in Søgaard and Olsen (2003). To homogenize the results we obtained a single displacement value, calculated as the average of the estimated deviation on the 10 bands. The algorithm has been rated on a dataset of test images:
   (a) note the barrel deformation introduced by the optics
   (b) the same images with digitally lens distortion correction
2. Proposed variant: apply a perspective correction to the images so that the rows are displayed parallel, then calculate the displacement as described in Søgaard and Olsen (2003) but on a single vertical histogram: in this way the information of the whole image is averaged in the histogram before calculating the phase. The algorithm has been rated on a dataset of test images:
   (a) note the barrel deformation introduced by the optics
   (b) the same images with digitally lens distortion correction
3. Proposed algorithm based on pattern matching. In this case it is not necessary to estimate the loss of precision due to lens distortion as it is already taken into account in the reference template. The algorithm has been rated on a dataset of test images:
   (a) note the barrel deformation introduced by the optics
   (b) limiting the analysis to only 10 or 20 equidistant lines of the test images

A deep learning approach was not chosen for the algorithm as it would require a large amount of data on which to train a model: as we will highlight later, creating a significantly large and varied dataset is not trivial.

2.1. Color segmentation adopted

For practical reasons, we have opted for a segmentation based on color (not just green crops), in order to recognize a greater variety of crops. Furthermore, compared to systems that seek a segmentation threshold based on the dual color distribution, we believe the solution of identifying the hue to be more robust for the confidence index in the presence of anomalous images: if a completely covered with plants or totally absent is taken coherent binary images are obtained (all detected or nothing) which maintain the desired sense of the confidence index, pushing it to low values.

Following the result of Li et al. (2020) that show the segmentation method based on the Lab color space allows obtaining, on average, good results over different crops and different lighting conditions, the strategy is considered convenient. In Lab color space 'L' channel represents the brightness, 'a' channel determines the green–magenta ratio while 'b' channel the yellow–blue ratio. Following Duarte-Correa et al. (2023) a portion of the area of interest in the image is selected on which the average value of channel 'a' and 'b' is calculated (target tint): the segmentation of the ROI takes place by selecting the pixels of the image which have values in the plane 'a,b' in a neighborhood of the target tint. For this purpose, the Euclidean distance between the points in the 'a,b' plane can be used to identify the neighborhood.

The evaluation of the segmentation method was empirical based on the comparison between the acquired images in the field (with different raspberry cameras (https://www.raspberrypi.com/products/), camera module v1, v2, v3, HqCam) and the area obtained with the segmentation. For a quantitative evaluation it is necessary to create, manually for a large dataset of collected images, the reference ground truth of the area covered by the plants. The experimental qualitative test has been verified that using a small distance threshold for 'a' channel allows to better discriminate plants from weeds while a greater distance threshold for 'b' channel allows to identify image areas with different light temperatures from the medium gray defined by AWB, i.e. identify in the same image ROI both in shady areas than sunny: in this way the set of points similar to the target tint in the 'ab' plane are contained in an ellipse instead of a circle. The method allows different degrees of freedom, to be analyzed experimentally as a function of the hardware and the crop.

Fig. 1 shows two examples of real images acquired in the field, partially shaded, which show the segmentation results discussed, superimposed in green.

2.2. Lens distortions

A lens, especially with wide-angle, introduces distortions to the acquired image. The phenomenon can be sufficiently described with radial distortion with respect to the distance from the center of the lens, using the first terms of the series expansion as a function. It is therefore possible to apply a computer vision algorithm (https://docs.opencv.org/2.4/doc/tutorials/calib3d/camera_calibration/camera_calibration.html, https://docs.wand-py.org/en/0.6.7/guide/distortion.html#barrel) at the captured video stream to correct its geometry, increasing the computational load. For specific needs there are also hardware solutions, i.e. manufacturers of lenses suitably designed and worked to reduce distortions.

In the algorithms that we are going to described and analyze below, the methods based on the histograms used the theoretical period of the seeding spacing calculated as a function of the installation geometry of the camera: if the distortion introduced by the lens becomes significant the histogram loses its periodicity, introducing error in the determination of the phase information as the algorithm is based on the assumption of analyzing a periodic signal with a known frequency. With the proposed algorithm based on template matching instead it is possible to precalculate the template with distortion: in this way the computational load on the video stream is not increase by a geometric transformation of the video stream and the measurement precision remains almost unchanged.

In the tests that we are going to propose, we will evaluate the measurement error with and without optical distortion correction assuming the use of a C-mount 6 mm lens with a FOV of 45° × 60°.

2.3. Analyzed algorithms

Considering having processed a binary image representative of the ROI of the video stream, below we describe and compare some methods of calculating the displacement of the camera with respect to the crop rows.

The methods are based on the assumptions described in the introduction, in particular: the sowing geometries are known, the camera installation geometry is fixed and defined, the camera moves parallel to the crop rows. In these hypotheses it is possible to calculate a priori the 2D perspective representation of the capture field, maintaining as the only degree of freedom the transverse displacement of the camera in the field, a value to be measured to implement the automatic inter-rows guidance.

2.3.1. Algorithm based on multiple vertical histograms

The first algorithm follows the work (Søgaard and Olsen, 2003). The image to be analyzed is divided into 10 horizontal bands, for each one the vertical histogram is calculated, i.e. the pixels values are added by columns (Fig. 2). Then each band is divided into its estimated sowing period in pixel, adding these last portions together we obtain the histogram representing the average period of the band. Finally taking as reference the maximum of the histogram, which corresponds to the thickening of the vegetation, the phase is calculated with respect to the period in order to estimate the position of the center row.

In the proposed work we introduce two variants: the first one we calculate the average phase from the phases of each band, from which we estimate the displacement. Second, we introduce a confidence index of the pattern, obtained as the difference of the maximum and minimum value of the average histogram normalized according to the maximum theoretical value of histogram (Eq. (1)); as for the phase, the overall index is the average of the indexes on the different bands. This index is normalized between 0 and 1, tends to one if it identifies an ideal sowing pattern (without weeds, camera parallel to the rows) and tends to 0 for patterns that are not representative of a row crop (it is not possible to distinguish the rows or with wrong direction).

$$Conf.Index_i = \frac{max(MeanHist_i) - min(MeanHist_i)}{yPixelCountHist_i * maxDataValue} \quad (1)$$

2.3.2. Algorithm based on single vertical histogram

Following the idea of Okamoto et al. (2001) a prospective transformation was applied to the acquired image in order to obtain the ground representation of the field; eliminating the perspective effect the rows appear parallel (first row of Fig. 3).

After that, following what was done with the previous algorithm, a single vertical histogram was computed and on it (second row of Fig. 3) and the phase and the confidence index were calculated.

In this case, since all lines of the transformed image have the same metric on the ground, it is possible to make a single average of the histogram and calculate the phase over a single period. In this way, from a computational point of view, only one phase search is performed.

In Fig. 3 we wanted to report the vertical histogram of the image without applying the perspective transformation (bottom left) to highlight a typical disturbed signal lacking information; a similar signal is

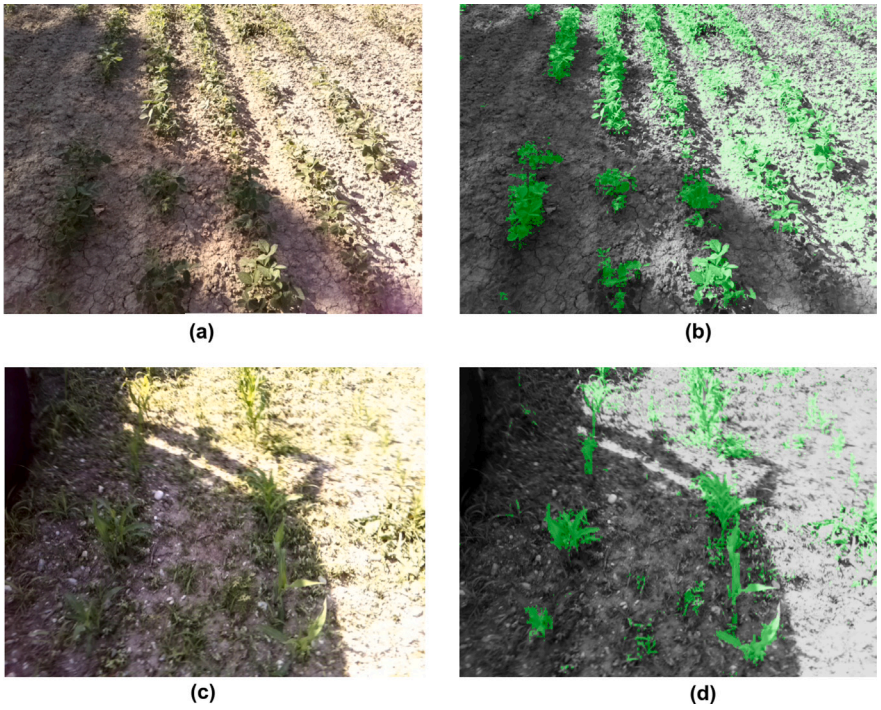

Fig. 1. Example of a field with soybean (a) and maize (c); highlighted in green over image their respective segmentations by Lab color space in (b) and (c).

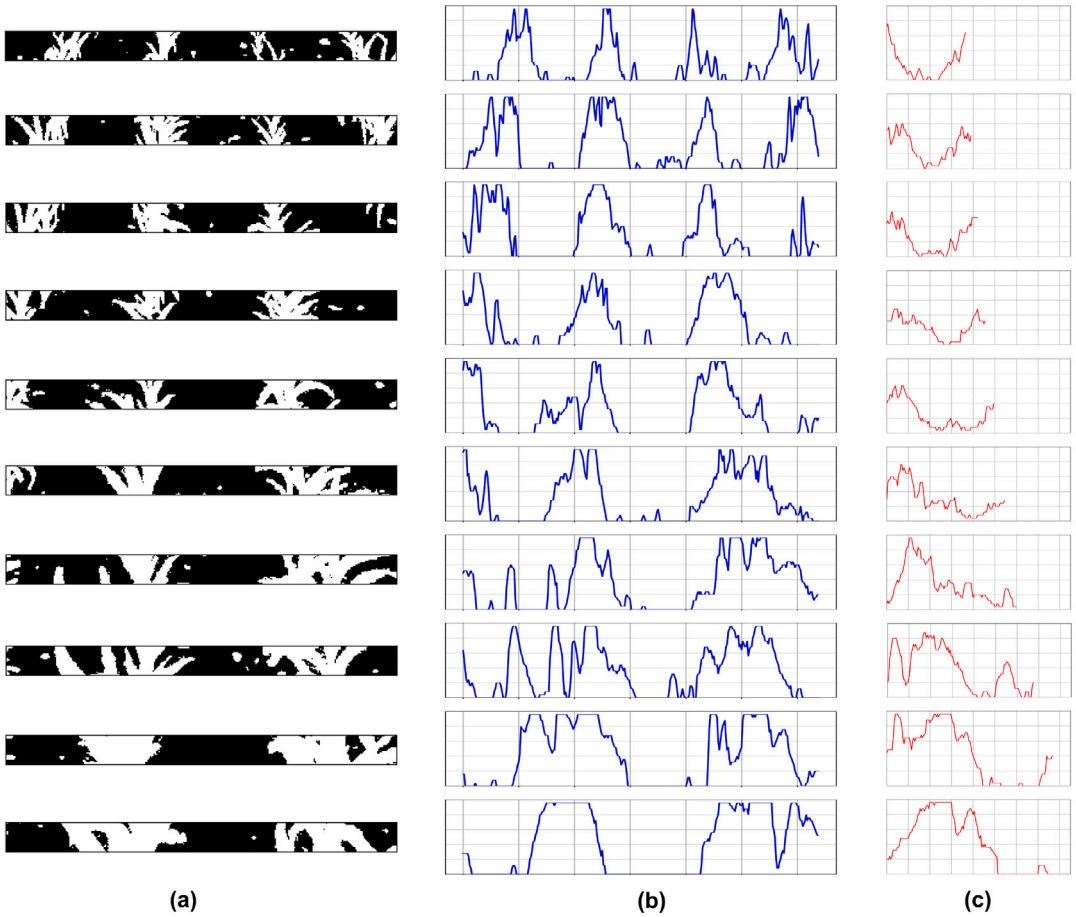

Fig. 2. Example of Søgaard and Olsen (2003) on a maize field. The binarized image is divided into 10 horizontal bands (a) and the vertical histogram is calculated for each (b); Considering previous knowledge of the period (expressed in pixels) for each band, the algorithm averages the histogram data over the presumed periods (c) to calculate the phase shift for each band.

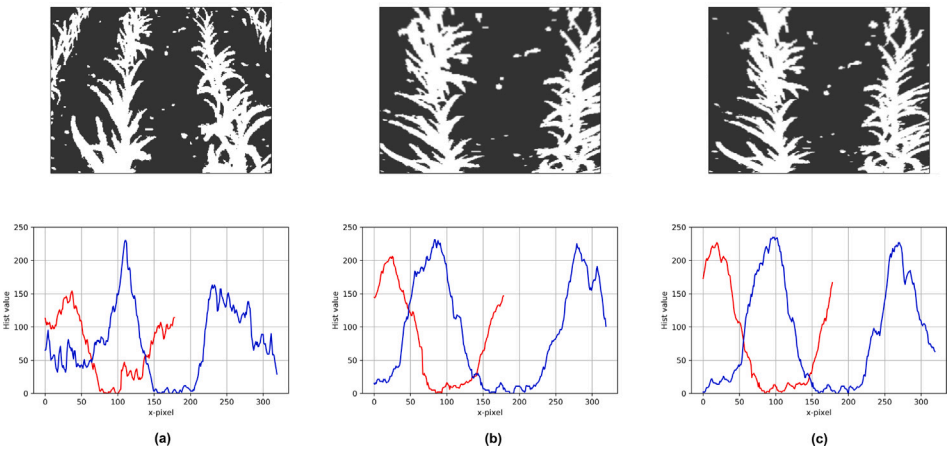

Fig. 3. Example of a maize field data: segmented image with respective vertical histogram (in blue) and average histogram on the knowing period (in red). (a) captured image, (b) image with only perspective transform, (c) image with both perspective transform and optical distortion correction.

obtained whenever the sowing pattern is strongly degraded or capture does not parallel with the crop rows. The next two histograms show how, with an appropriate perspective transformation, a signal with useful information emerges, with well-defined peaks in correspondence of the crops rows: if the correction of optical distortion is also applied (below right) the histogram improves, becoming less scattered.

2.3.3. Algorithm based on pattern matching

The algorithm that we are going to propose, as anticipated, is based on the creation of a template known the geometry of sowing and of the camera through which to estimate the deviation of the crop rows respect to a reference. It is required that the following are known in advance: height and angle of inclination respect to the ground of the camera, vertical and horizontal FOV of the optics, coefficients of distortion of the lens; for the crop, the sowing pitch of the rows, the average height and width of the plants. For the specific weeding application to which the project is addressed, these data are certainly known a priori and are invariant during work; the same data are also necessary for the calculation of the theoretical period in the two previous algorithms. With these information, by means of trigonometric calculations and scaling the metric quantities in pixels, it is possible to identify the row centers on the ground and their perspective vanishing point, that is centered in width as the rows are travel lengthwise. With the dimensional information of the plant, as a function of their state of growth, the volume occupied around the sowing line is estimated, again using a perspective representation. With these data the reference image – the template – is created (Fig. 4); finally it is distorted radially to simulate the artifact of the lens. Obviously, similarly to what happens with the previous algorithms in which the period to be searched is defined, the template refers to the ideal case with a camera that capture crop rows aligned to the vertical and perfectly straight. Slight rotations or curvatures of the rows are tolerated and will lower the maximum correlation value and confidence index.

The estimation of the displacement between camera and field takes place by searching for the displacement of the template that maximizes the correlation with the video stream. Since the field is travel parallel to the rows, the movement of the cultivator can only perpendicular to the crop rows: the search for the seeding template takes place only in one dimension of the image, along the x axis. Since moving the shot changes the position at the base of the rows while the vanishing point remains unchanged, it is necessary to generate a template for each movement and not simply move it spatially in the 1-D convolution operation. For these reasons, a set of templates is created beforehand for a certain number of movements. Having fixed a reference configuration, for example in which the rows assume vertical symmetry i.e. the camera is centered in the inter-row space, templates can be generated for displacements of one centimeter up to displacements of 20 cm, an area in which maximum precision is desired, to switch to milder movements to cover greater distances, up to covering the periodicity of the crop rows. For this algorithm, the confidence index was calculated as the ratio between the difference and the sum of the maximum and minimum correlation value over the shifts (Eq. (2)), always to obtain an index in the 0–1 range with values tending to one for well recognizable fields.

$$Conf.Index = \frac{max(Correlation) - min(Correlation)}{max(Correlation) + min(Correlation)} \tag{2}$$

2.3.4. Proposed simplification

Having to work with a binary input image as it is segmented by color, it is advisable to create templates with binary values and define the multiplicative part of convolution with binary operator in order to speed up the calculation, especially if on dedicated hardware. Using the boolean algebra, the multiplicative part of correlation operation corresponds to an XNOR. Since several templates have already been calculated at the same resolution as the image to be analyzed, the XNOR operation is performed pixel by pixel between each template and the image; the similarity score (the correlation value) for each template is obtained by counting how many pixels have XNOR to one (Fig. 5) that is, with an addition operation. So the in-phase template will have the maximum correlation index while it will be minimal for the quadrature template. Obviously, using the primitive boolean operator XOR the phase estimate passes from the search for the minimum of the XOR sums.

To limit the number of operations, which tend to increase using a wide number of templates to have better measurement resolution, it is proposed to limit the analysis to a few tens of equally spaced lines of the image, since the vertical information is redundant (Fig. 6). This expedient leads to a lesser loss of measurement precision if a median filtering followed by a dilatation is applied to the captured image before to be processed by lines so as to preserve, in the few lines analyzed, information on the surrounding discarded pixels (Fig. 7); obviously, still with a view to reducing the computational load, filtering can be limited only around the lines that will be analyzed.

The Fig. 8 shows a flowchart of the proposed algorithm with the simplifications described. The algorithm is applied to each frame acquired by the camera so its execution time must be less than the frame rate of the video stream. The mechanical movement control system of the weeder will receive data at the frame rate cadence: it will have to interpret the confidence value in order to block the movement or move the weeder elements according to the deviation measurement received.

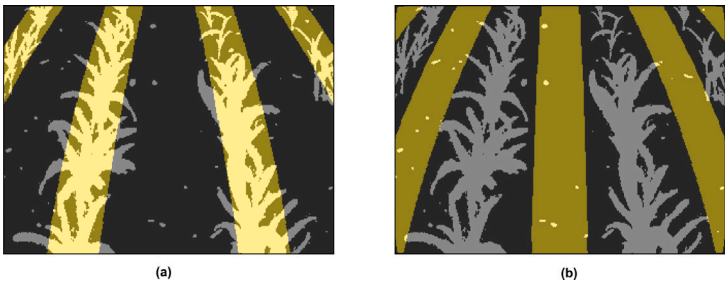

Fig. 4. Binary image to be analyzed (in black and white) with the created reference template superimposed in yellow. (a) the crop rows are in phase with the chosen template, (b) the crop rows are out of phase with the chosen template (reference template shifted by 180° i.e. by half a row-space).

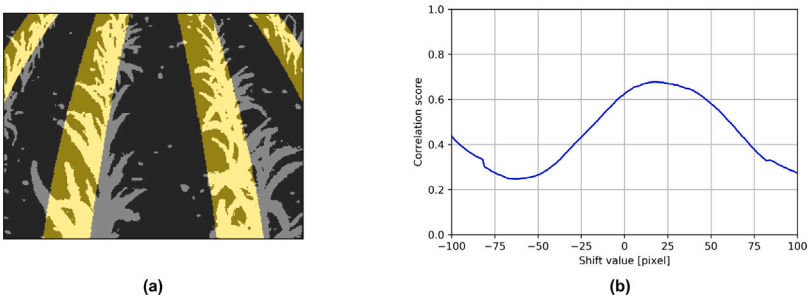

Fig. 5. (a) example of a maize field with a reference mask superimposed in yellow; the crop rows was located slightly to right respect the reference. (b) graph of the normalized sums of the XNOR calculated for different displacements of the template; using the XNOR operator, the estimated displacement corresponds to the maximum sum value: the peak of the graph is shifted to the right like the crop rows.

Fig. 6. Binary image to be analyzed (in black and white) with the created reference template superimposed in yellow. (a) computing only 20 lines of image 5, (b) analyzing only 10 lines.

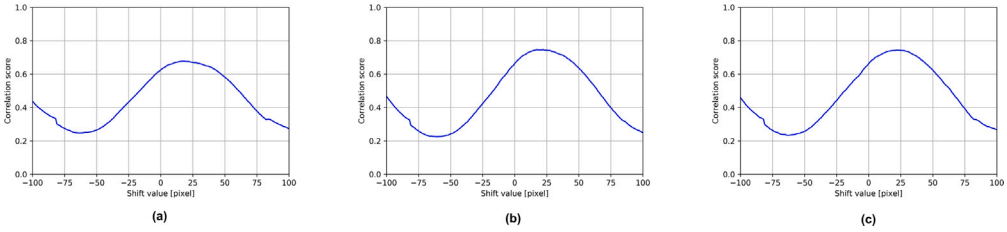

Fig. 7. Graph of the normalized sums of the XNOR calculated for different displacements of the template. (a) computing on the whole image, (b) analyzing only 20 lines of image, (c) analyzing only 10 lines.

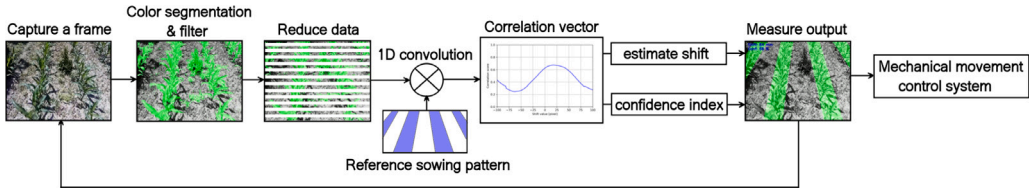

Fig. 8. Flowchart of the proposed algorithm.

2.4. Dataset

In order to compare and validate the vision algorithms considered and proposed, it is necessary to have a dataset of images:

- which allows to numerically analyze and compare different algorithms referring to the same data;
- which allows to tune an algorithm under development;
- that has significant variance to obtain a measurement statistic.

For this purpose a shared dataset of images of a crop rows fields with similar characteristics to the cultivator problem under study was not found.

Example of plant templates used to generate maize field models

Example of plant templates used to generate soybean field models

Fig. 9. Example of plant templates used to generate the dataset.

One possibility is to create it through field acquisitions with controlled movement of a camera. This strategy, however, brings some problems:

- it is difficult to automatically generate the corresponding ground truth;
- it is difficult to find or create a sown field with certain characteristics, such as a certain predetermined number of missing plants;
- the positioning errors of the camera respect to the field will enter in the validation measurement chain;
- in the validation measurement chain also enter the capture system's artifact errors i.e. the sensor color profile, rolling shutter artifacts, white balance, lens distortion.

The last point can be corrected with image pre-processing, techniques closely related to the vision system used. Furthermore, cultivating a field to create a dataset presents practical difficulties related above all to seasonality and crop development times.

To achieve uniformity in the comparison of vision algorithms and to be able to explore specific aspects, it was therefore decided to create a dataset of synthetic images with the desired characteristics. A **digital twin** of a sown field was then created, allowing its parameterization during generation. This two-dimensional dataset is a valid representation of the video stream that the algorithms will process.

In particular, an algorithm capable of generating synthetic images with a certain degree of diversity was written by parameterizing some geometric characteristics of the problem. The parameters of the camera are defined such as FoV, height and angle from the ground, while for the field are defined the row, inter-row spacing and height of the plants, percentage of missing and double plants, having first generated a certain number of synthetic plants for each variety (Fig. 9 shows some examples, obtained by manually segmenting BW photos with GIMP); finally the algorithm is capable of generating images at different displacements of the camera with respect to the inter-row center. All algorithms developed and analyzed for this work were written by the authors in Python, using the Numpy and OpenCV libraries.

Similarly to what was done to create the sowing template, the grid on the ground of the base of the plants is calculated trigonometrically to then place images of plants (about thirty different images, with slight random rotation, random axial flip, color alteration and dither respect theoretical position) suitably scaled in perspective height. The digital twin created is a two-dimensional dataset (image), a valid representation of the video stream, captured with a camera subjected to precise constraints, on which the algorithms operate.

Two datasets were created one with maize plants on a brown soil, about 50 cm tall sown at 20 cm with a row spacing of 75 cm, and one with soybean plants, about 30 cm tall sown at 4 cm with a 45 cm row spacing. Realistically, we plan to use a camera 1.1 m high from the ground inclined towards the horizon of 45°, with FOV 45° × 60°.

These datasets were generated with displacements of up to ±60% of the row spacing at 1 cm steps, 50 images are generated for each displacement by varying: the positioning of the plants, adding a noise of ±5° of the vanishing point, adding interrow weeds randomly (with color similar to plants so as to disturb the segmentation), randomly removing some plants and adding partial shading or light flare to the image. For both crops, combinations were generated by varying the cover of weeds (in terms of 5%, 20%, 40%, 60%, see Figs. 10, 11) and the number of missing plants (in the order of 5%, 40%, 60%, see Fig. 12).

In the examples shown in the preceding figures, the greens of plants and weeds are diversified for clarity. To analyze the different algorithms, the datasets were segmented according to color, in order to obtain a binary image: to maintain the sense of the different tests, i.e. verify the robustness of the algorithms as weeds grow, the segmentation was conducted by greatly expanding the colors to recognize, accepting almost all green pixels.

Fig. 13 intended to show the similarity of the digital model compared to some photos taken in the field with the same settings, thus visually validating its representativeness.

2.5. Measures

The images of the dataset have been segmented according to the described Lab method, using a wide color threshold so that all weeds appear together with the plants in the binary images: the aim is to evaluate the sensitivity of the algorithms to different levels of disturbance. Similarly, simulating missing plants on lightly infested land serves to evaluate the robustness of the algorithms in critical situations. For a more realistic simulation it was preferred to generate a color dataset with the various noises explained and segmented it rather than use ideal binary image.

A first evaluation of the algorithms was conducted by applying them to the entire datasets: the 50 images were used to calculate the mean and variance of the measure for each shifts. The graph of these values shows the linearity and measurement dispersion of the methods, highlighting when the method begins to be affected by the periodicity of crop rows.

Since limited displacements around the reference are of interest for automatic guidance, the subsequent characterization measurements of the algorithms were carried out for displacements within ±30% of the row spacing, i.e. approximately within ±20 cm around the row plants. For this type of measurement, the 50 images per displacement were

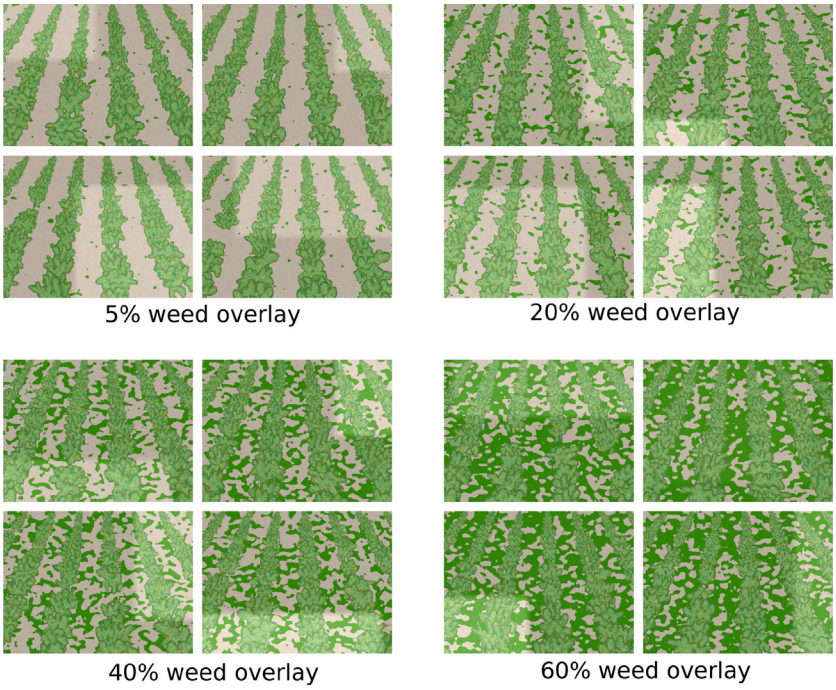

Fig. 10. Example of soybean dataset with different levels of weed coverage.

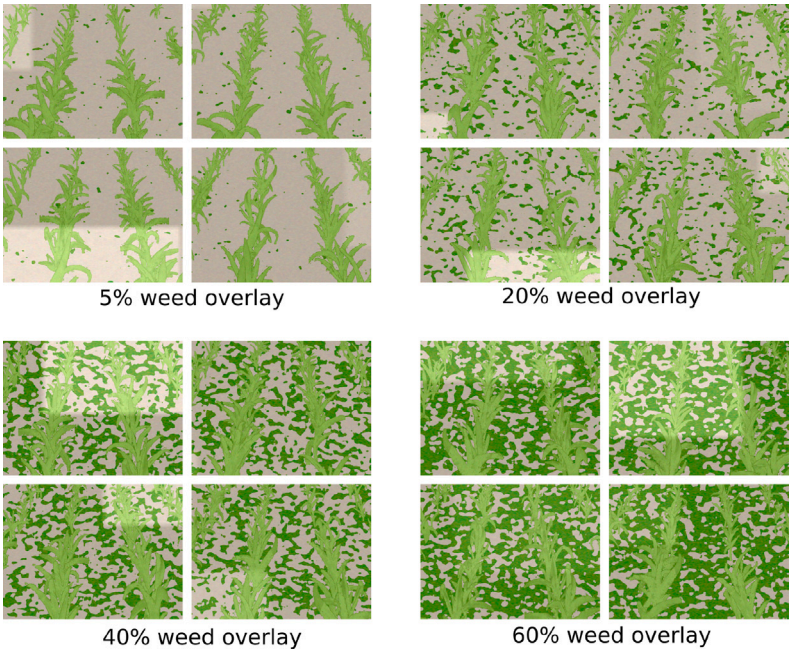

Fig. 11. Example of maize dataset with different levels of weed coverage.

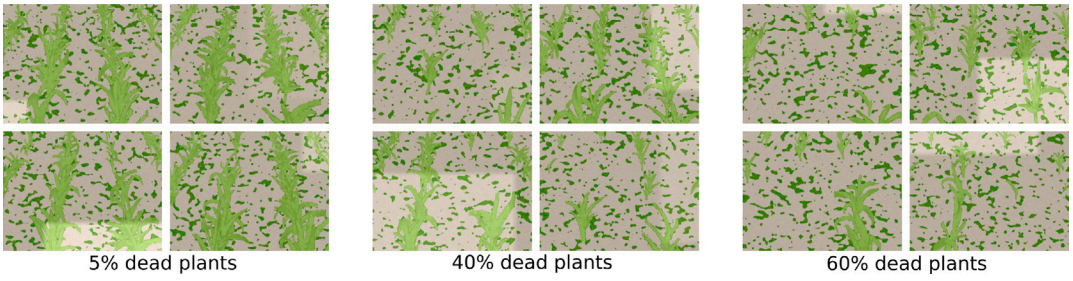

Fig. 12. Example of maize dataset with different levels of dead plants.

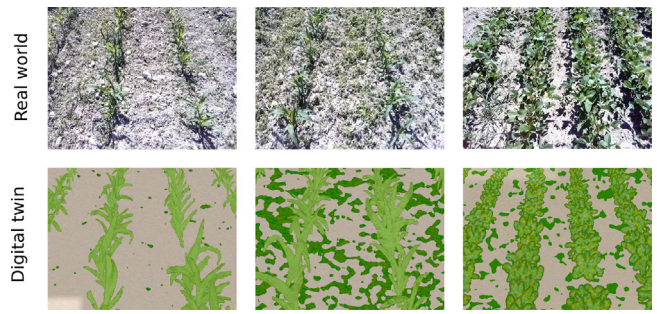

Fig. 13. Examples of maize and soybean field compared with their digital twin.

used to obtain the statistical data of variance; to define the precision of the system, double the average value of standard deviations over the analyzed displacement interval is considered, therefore corresponding to the dispersion of 95% of the measurements. For comparison we do not consider interest the, limited, mean value as it represents the offset of the measurement system. For these tests, in addition to the precision value, the value of the index r^2 (Devore, 2011) is reported, representing the linearity of the measurement.

This last test was conducted on the two datasets, maize and soybeans. The two algorithms based on the vertical histograms were evaluated on the dataset images both without and with optical distortion in order to evaluate how much degrades the measurement accuracy when saving the computational load of correcting the distortion. For the proposed algorithm, based on template matching, the precision is evaluated in the case of processing the full image or only 20 or 10 of its lines; as mentioned, since the template is already predistorted, it makes no sense to evaluate the effect of optical distortion. Finally, the result of the proposed algorithm will be shown on real images captured in the field, in order to highlight the consistency of the numerical validation carried out on the synthetic dataset.

3. Results and discussion

The results proposed below were obtained by processing images at a resolution of 240 × 320 pixels. This resolution was chosen because it offers a good compromise between measurement accuracy and computational load.

3.1. Measurement feature

The results of the first type of tests, where the response of algorithms are evaluated for a wide range of displacements, are reported in Figs. 14, 15, 16. For this test, the database generated with the lowest number of weeds and missing plants was used, in order to globally evaluate a non-borderline case. For the algorithms based on the analysis of vertical histograms, the figures on the left show the response over the shifts using images without lens distortions while on the right the response of algorithms if distorted images are used.

Instead, for the proposed algorithm based on template matching, the response are reported if the whole image is processed (on left), only 20 equidistant lines (in center) or only 10 lines (on right).

This test highlight the linear response of the different algorithms for an extended range of displacements from the reference; only at the limits of the range considered, correctly, the measurement error increases due to the periodicity of the sowing pattern. Furthermore these test highlight, for the algorithms based on histograms, how the use of an image corrected by optical distortion improves the measurement limiting their dispersion while, for the template matching method, the measurement does not suffer great dispersion in precision when analyzing only a limited set of lines of the image.

3.2. Accuracy

The results of the second type of test, in which the measurement accuracy is evaluated for limited movements around the central row, are reported in the Tables 1, 2, 3, 4 in terms of measurement precision, considering double the average value of standard deviations, and measurement linearity, using the index r^2. For better overall readability, the data are displayed graphically in the Figs. 17–24.

These tests want to highlight the measurement accuracy for movements around the target position; if the deviation is high, as the tests in the previous section have shown, the systems are still able to provide a correction value for the position, albeit with less precision.

In particular these tests highlight two borderline cases that can be encountered in the real problem of driving cultivator, both of which make it difficult to recognize the sowing pattern and therefore make an incorrect estimate of its displacement: an increase in weeds or an excess of missing plants.

The results show, for both crops and for both field anomalies, a trend consistent with expectations, i.e that the correction of the optical distortion leads to a more accurate measurement while limiting the pattern matching analysis to a few lines degrades the accuracy.

It is interesting to highlight how working on images with optical distortion weighs heavily on the analysis method with multiple vertical histograms, since the phase information obtained in the more peripheral bands is strongly wrong: as the reference paper suggests, the maximum accuracy is obtained in the central band, where the optical distortion is less. By first correcting the perspective instead and calculating a single histogram there are two advantages: the distortions at the high edges are eliminated in the affine transformation, the distortions that remain have little weight in the single vertical average.

The proposed pattern matching method shows greater measurement precision in all tests, with acceptable values even in highly borderline cases. Even in the case of limiting the analysis to only 10 of the 240 lines of the images, precision indexes are comparable to the methods taken as a reference. It has been estimated that filtering the segmented images with a median filter (in the order of 7 × 7 pixel on 320 × 240 pixel image size) does not improve the reported indices except for the method based on pattern matching when the analysis is limited to a few rows, in which case the use of the filter improves the precision by up to 5 mm: this is due to the effect of the median filter, which also carries information from the surrounding lines which are omitted in the calculation.

On the other hand, if the resolution of the image is increased from 240 × 320 to 480 × 640 or 600 × 800 pixel, there are no significant increases in the two precision indexes considered, so we do not believe that increase the computational load.

3.3. Confidence index

The results shown in the Tables 5, 6, 7, 8 refer to the same type of test as in the previous section: they complete the analysis of the proposed measurement algorithms, highlighting when they are no longer able to reliably recognize the seeding pattern. For better overall readability, the data are displayed graphically in the Figs. 25–28.

The aim is to identify a unique threshold for the confidence index such as to switch off the position control system when the deviation measurement loses its meaning.

Similarly to precision, the confidence index also decreases as the non-ideal nature of the analyzed image increases. In particular, for methods based on vertical histograms there is a rapid drop in the confidence index together with an increase in the measurement error, making it difficult to find a threshold value that also includes cases with disturbances.

For the proposed method based on pattern matching, on the other hand, against a drop of tens of percentage points in the confidence index, proving that the sowing pattern captured is irregular, it is still

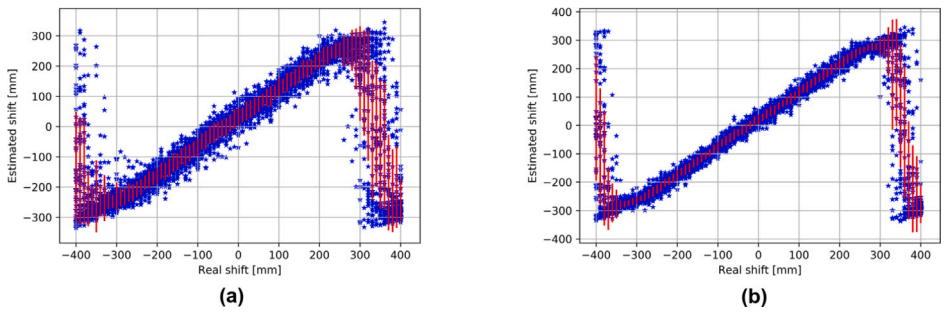

Fig. 14. Algorithm based on multiple vertical histograms: response over a wide range of displacements. Blue stars: scatter measures; red bars: std over 50 images. Test using: (a) images without correction of lens distortions, (b) corrected images.

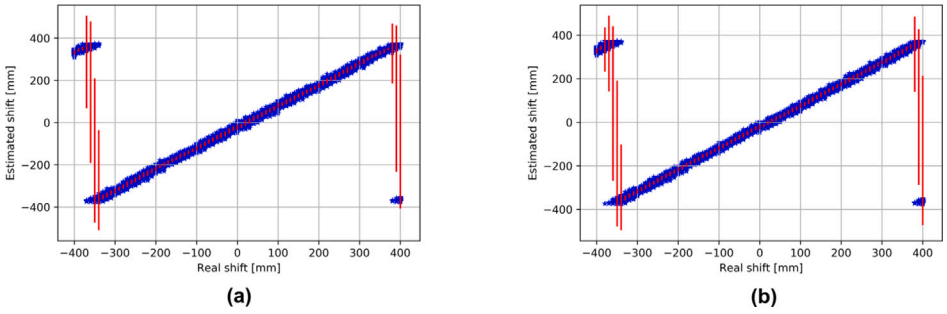

Fig. 15. Algorithm based on a single vertical histogram: response over a wide range of displacements. Blue stars: scatter measures; red bars: std over 50 images. Test using: (a) images without correction of lens distortions on left, (b) corrected images on right.

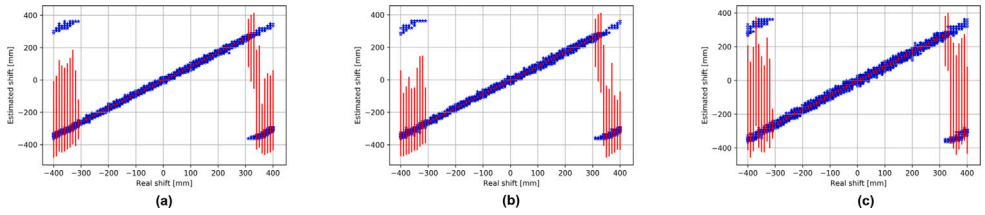

Fig. 16. Algorithm based on pattern matching: response over a wide range of displacements. Blue stars: scatter measures; red bars: std over 50 images. (a) process all lines of image on left, (b) only 20 lines in middle, (c) only 10 lines in right.

Table 1
Comparison of accuracy on maize dataset by varying the level of weeds.

Algorithm	Variant	5% weed		20% weed		40% weed		60% weed	
		precis.	r^2	precis.	r^2	precis.	r^2	precis.	r^2
Hists	With Dist.	57 mm	0.94873	57 mm	0.94440	68 mm	0.91358	74 mm	0.87669
Hists	No Dist.	32 mm	0.98337	35 mm	0.98062	45 mm	0.96834	53 mm	0.95465
Hist	With Dist.	22 mm	0.99277	28 mm	0.98796	40 mm	0.97599	56 mm	0.95586
Hist	No Dist.	20 mm	0.99365	26 mm	0.98941	35 mm	0.98051	49 mm	0.96411
Pattern	All Lines	18 mm	0.99370	20 mm	0.99187	25 mm	0.98776	30 mm	0.98308
Pattern	20 Lines	19 mm	0.99260	23 mm	0.98960	28 mm	0.98482	38 mm	0.97505
Pattern	10 Lines	26 mm	0.98637	30 mm	0.98181	40 mm	0.96936	50 mm	0.95404

Table 2
Comparison of accuracy on maize dataset by varying the level of missing plants.

Algorithm	Variant	5% missing		40% missing		60% missing	
		precis.	r^2	precis.	r^2	precis.	r^2
Hists	With Dist.	57 mm	0.94444	89 mm	0.86888	112 mm	0.76167
Hists	No Dist.	35 mm	0.98062	55 mm	0.95192	77 mm	0.89855
Hist	With Dist.	28 mm	0.98796	43 mm	0.97230	70 mm	0.93146
Hist	No Dist.	26 mm	0.98941	33 mm	0.98307	50 mm	0.96774
Pattern	All Lines	20 mm	0.99187	27 mm	0.98470	39 mm	0.96795
Pattern	20 Lines	23 mm	0.98960	31 mm	0.98142	41 mm	0.96615
Pattern	10 Lines	30 mm	0.98181	41 mm	0.96597	56 mm	0.93459

Table 3
Comparison of accuracy on soybean dataset by varying the level of weeds.

Algorithm	Variant	5% weed		20% weed		40% weed		60% weed	
		precis.	r^2	precis.	r^2	precis.	r^2	precis.	r^2
Hists	With Dist.	94 mm	0.87739	92 mm	0.95203	83 mm	0.82214	92 mm	0.82077
Hists	No Dist.	15 mm	0.99095	22 mm	0.98014	27 mm	0.96887	38 mm	0.93465
Hist	With Dist.	17 mm	0.98889	17 mm	0.98862	19 mm	0.98626	27 mm	0.97411
Hist	No Dist.	16 mm	0.99132	17 mm	0.98956	18 mm	0.98778	22 mm	0.98123
Pattern	All Lines	9 mm	0.99640	10 mm	0.99548	12 mm	0.99304	20 mm	0.98445
Pattern	20 Lines	11 mm	0.99489	12 mm	0.99363	15 mm	0.99113	22 mm	0.98017
Pattern	10 Lines	13 mm	0.99233	17 mm	0.98808	23 mm	0.97901	35 mm	0.95429

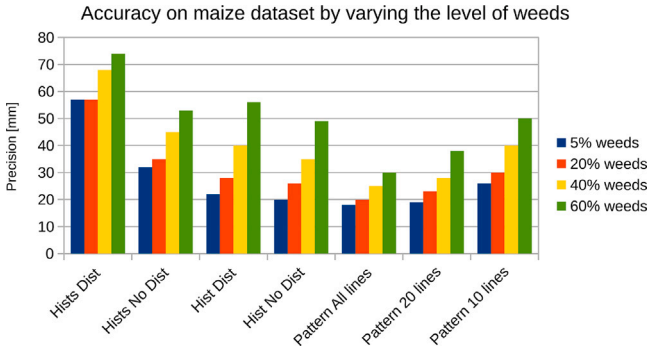

Fig. 17. Comparison of accuracy, in term of average precision, on maize dataset by varying the level of weeds.

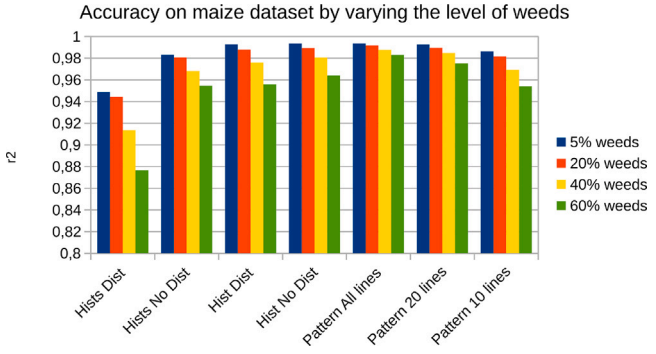

Fig. 18. Comparison of accuracy, using the r^2 index, on maize dataset by varying the level of weeds.

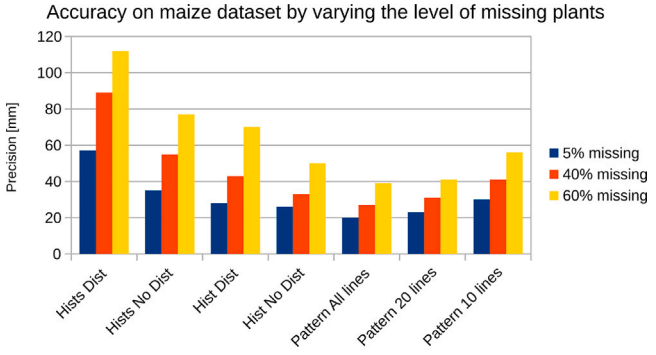

Fig. 19. Comparison of accuracy, in term of average precision, on maize dataset by varying the level of missing plants.

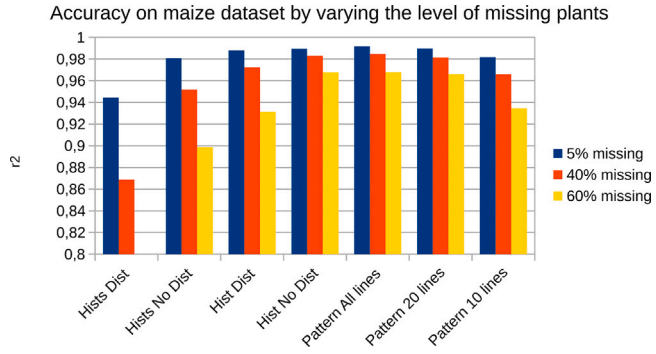

Fig. 20. Comparison of accuracy, using the r^2 index, on maize dataset by varying the level of missing plants.

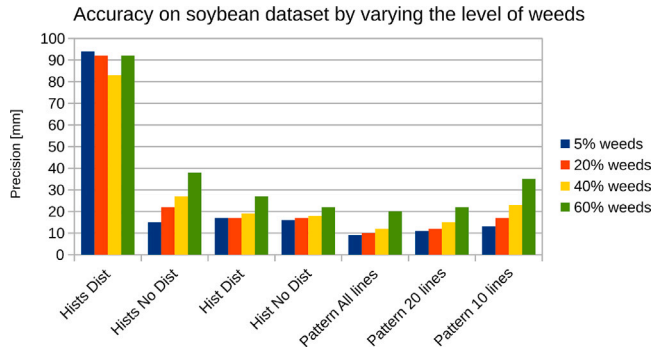

Fig. 21. Comparison of accuracy, in term of average precision, on soybean dataset by varying the level of weeds.

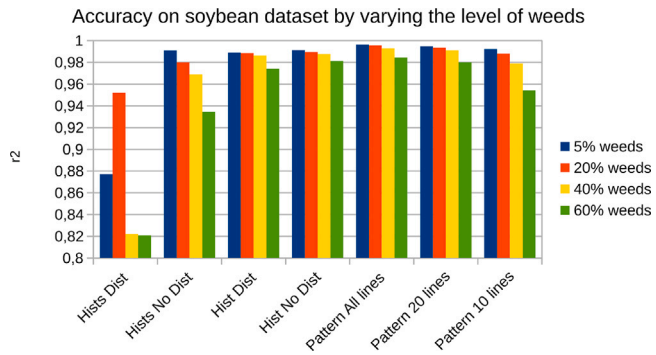

Fig. 22. Comparison of accuracy, using the r^2 index, on soybean dataset by varying the level of weeds.

possible to maintain good precision on the deviation measurement. Also for this index, limiting the analysis to a few lines of the image maintains the original information content. From the proposed tests, the new method is therefore more robust in simulated borderline cases.

3.4. Time analysis

The proposed comparative tests were analyzed on a PC with i5-4440 4 × 3.10 GHz processor, described in Python language using

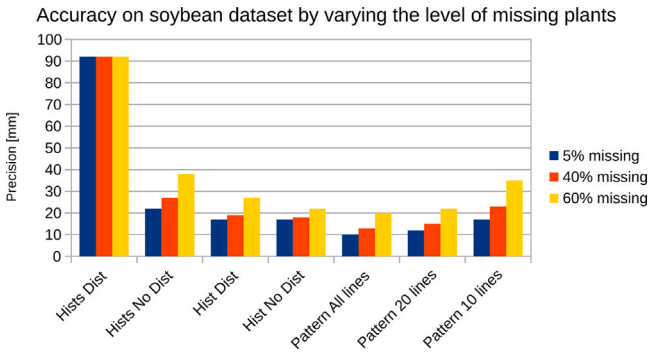

Fig. 23. Comparison of accuracy, in term of average precision, on soybean dataset by varying the level of missing plants.

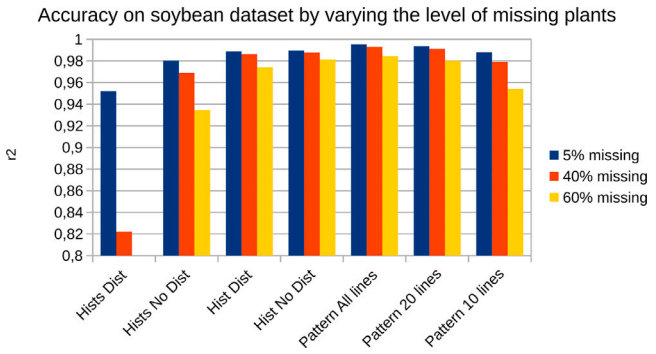

Fig. 24. Comparison of accuracy, using the r^2 index, on soybean dataset by varying the level of missing plants.

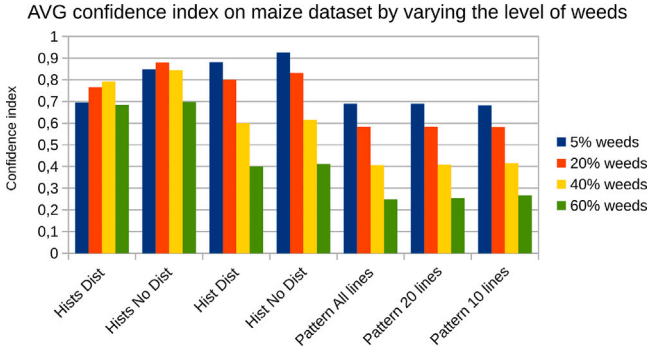

Fig. 25. Comparison of the average confidence index on maize dataset by varying the level of weeds.

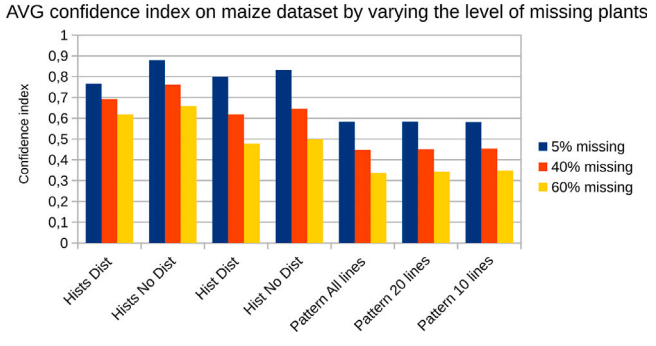

Fig. 26. Comparison of the average confidence index on maize dataset by varying the level of missing plant.

Table 4
Comparison of accuracy on soybean dataset by varying the level of missing plants.

Algorithm	Variant	5% missing		40% missing		60% missing	
		precis.	r^2	precis.	r^2	precis.	r^2
Hists	With Dist.	92 mm	0.95203	92 mm	0.82214	92 mm	0.66026
Hists	No Dist.	22 mm	0.98014	27 mm	0.96887	38 mm	0.93465
Hist	With Dist.	17 mm	0.98862	19 mm	0.98626	27 mm	0.97411
Hist	No Dist.	17 mm	0.98956	18 mm	0.98778	22 mm	0.98132
Pattern	All Lines	10 mm	0.99548	13 mm	0.99304	20 mm	0.98445
Pattern	20 Lines	12 mm	0.99363	15 mm	0.99113	22 mm	0.98017
Pattern	10 Lines	17 mm	0.98808	23 mm	0.97901	35 mm	0.95429

Table 5
Comparison of the average confidence index on maize dataset by varying the level of weeds.

Algorithm	Variant	5% weed	20% weed	40% weed	60% weed
Hists	With Dist.	0.6955	0.7666	0.7914	0.6837
Hists	No Dist.	0.8486	0.8800	0.8438	0.6983
Hist	With Dist.	0.8811	0.8003	0.5996	0.4009
Hist	No Dist.	0.9267	0.8320	0.6145	0.4114
Pattern	All Lines	0.6897	0.5830	0.4061	0.2490
Pattern	20 Lines	0.6897	0.5832	0.4095	0.2541
Pattern	10 Lines	0.6818	0.5814	0.4154	0.2678

Table 6
Comparison of the average confidence index on maize dataset by varying the level of missing plant.

Algorithm	Variant	5% missing	40% missing	60% missing
Hists	With Dist.	0.7666	0.6920	0.6190
Hists	No Dist.	0.8800	0.7618	0.6590
Hist	With Dist.	0.8003	0.6192	0.4778
Hist	No Dist.	0.8320	0.6451	0.4986
Pattern	All Lines	0.5830	0.4480	0.3376
Pattern	20 Lines	0.5832	0.4518	0.3433
Pattern	10 Lines	0.5814	0.4541	0.3486

Table 7
Comparison of the average confidence index on soybean dataset by varying the level of weeds.

Algorithm	Variant	5% weed	20% weed	40% weed	60% weed
Hists	With Dist.	0.5481	0.5530	0.5236	0.4619
Hists	No Dist.	0.5974	0.5958	0.5406	0.4605
Hist	With Dist.	0.8976	0.7393	0.5065	0.3016
Hist	No Dist.	0.9581	0.8021	0.5571	0.3342
Pattern	All Lines	0.6892	0.6101	0.4555	0.2494
Pattern	20 Lines	0.6742	0.5978	0.4488	0.2506
Pattern	10 Lines	0.6613	0.5859	0.4404	0.2499

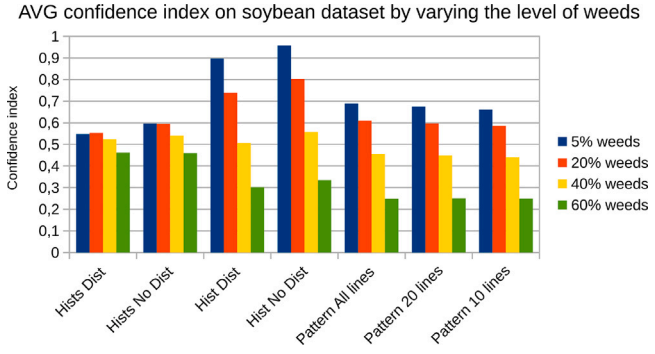

Fig. 27. Comparison of the average confidence index on soybean dataset by varying the level of weeds.

Table 8
Comparison of the average confidence index on soybean dataset by varying the level of missing plant.

Algorithm	Variant	5% missing	40% missing	60% missing
Hists	With Dist.	0.7666	0.6920	0.6190
Hists	No Dist.	0.8800	0.7618	0.6590
Hist	With Dist.	0.8003	0.6192	0.4778
Hist	No Dist.	0.8320	0.6451	0.4986
Pattern	All Lines	0.5830	0.4480	0.3376
Pattern	20 Lines	0.5832	0.4518	0.3433
Pattern	10 Lines	0.5814	0.4541	0.3486

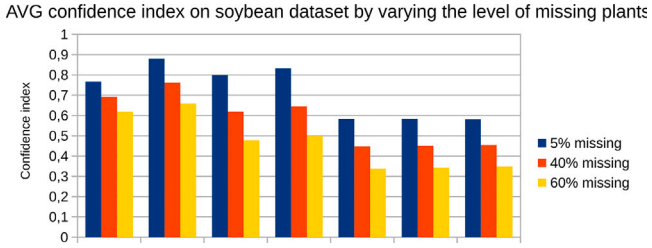

Fig. 28. Comparison of the average confidence index on soybean dataset by varying the level of missing plant.

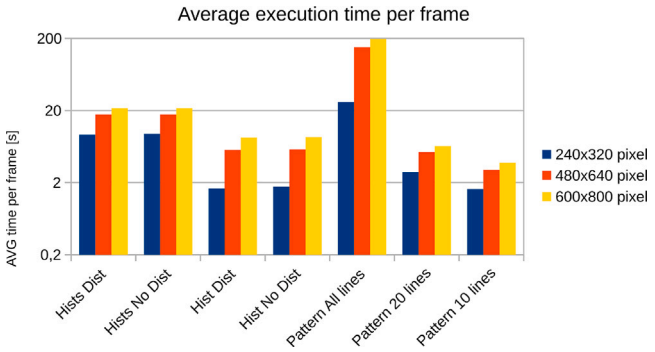

Fig. 29. Comparison of the average execution time per frame of the algorithms at different resolutions of the input image.

OpenCV and Numpy library. Table 9 shows the average execution times per frame of the described algorithms, time including color segmentation and displacement estimation; the data is displayed graphically in the figures the data is displayed graphically in Figs. 29 and 30. Together with the simulation execution time the table shows the estimates computational complexity for each image, in terms of floating point operations per frame (FLOPF), to give a hardware independent comparison.

The times measured in the simulations, however, do not follow the ratios indicated by the FLOPF estimates: this is due to the code optimizations with which the used libraries have been developed for the specific hardware architecture.

A good saving of time is observed in analyzing a single vertical histogram instead of analyzing several bands of the image, highlighting, on a microprocessor system, how the perspective transformation weight much less time consuming than calculating the phase several times on the different bands. Even the optical distortion correction weighs little in the overall time, allowing an increase in measure precision.

The pattern matching method, on the other hand, is very time-consuming if applied to the entire image; however, the various tests proposed show how limiting the pattern analysis to only 20 lines of the image leads to a considerable speedup of the code, both in terms of time and computation, decreasing the measurement precision by only a few mm. In conclusion, the proposed algorithm allows to obtain

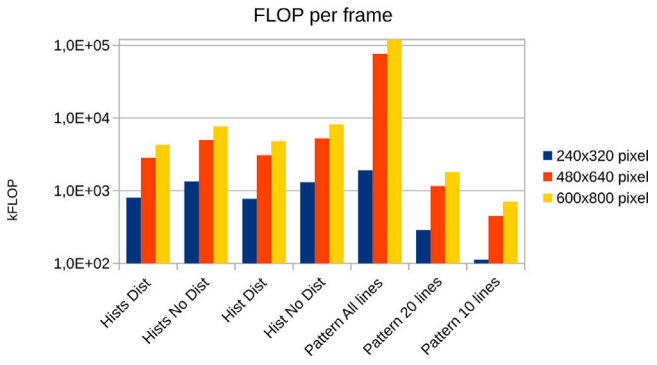

Fig. 30. Comparison of FLOP per frame of the algorithms at different resolutions of the input image.

greater precision than the two methods compared with execution times comparable to the fastest algorithm. Surely if the pattern matching algorithm were implemented on dedicated hardware, the possibility of performing XOR at the bit level would lead to a further speedup.

Versions of the 20-line pattern matching algorithm and the single vertical histogram algorithm were tested on a Raspberry PI4 SBC (Single Board Computer) with a CSI RaspyCam, showing real-time capability up to 15 fps on a 240 × 320 pixel video stream.

3.5. Proposed algorithm

A qualitative validation of the presented algorithm is proposed by applying it to real images captured in the field, using a camera setup identical to the parameterization used to generate the digital dataset. For this analysis, videos were saved and subsequently processed on the PC in order to evaluate the tuning of the free parameters such as the color segmentation threshold, resolution and lines to be analyzed. Fig. 31 shows some significant frames of these videos taken in the field, highlighting on the side the position identified for the sowing pattern and the confidence index.

The examples (a–d) demonstrate accurate row identification: the overlay of the pattern chosen by the algorithm shows its validity in estimating the displacement of the crop rows respect to the camera with good tolerance to weeds and gaps. Furthermore, in these cases the algorithm calculates a high confidence index which reflects the quality of the sowing.

Conversely, in example (e) the rows are rotated while in example (f) the sowing is crossed: in these anomalous use cases the algorithm always estimates a deviation of the pattern, since it is related to the maximum correlation value, but calculates a low confidence index, suggesting that these measurements should be rejected. A low confidence index in these cases indicates that the tractor is moving in a direction unsuitable for effective cultivation, potentially compromising the accuracy of soil tilling: the controller of the weeding element's movement must therefore take into account both measures of the algorithm.

Under particular conditions such as crossing rows or anomalous shadowing patterns, the confidence index exhibits a clear reduction. Such behavior can be usefully integrated in field operations, producing a feedback for the farmer, whose attention might be recalled in the case of field conditions characterized by a loss of orientation of crop rows.

The compromise to be found is to identify the shortest execution time while maintaining resolution in the measurement. Having a high frame rate allows higher tractor speeds as it guarantees capturing images without blur and minimizing the space on the ground between successive measurements, thus providing better and continuous data to the position controller.

Table 9
Comparison of the average execution time and FLOP per frame of the algorithms at different resolutions of the input image.

Algorithm	Variant	240 × 320 pixel		480 × 640 pixel		600 × 800 pixel	
		Time	FLOP	Time	FLOP	Time	FLOP
Hists	With Dist.	9.25 ms	802 k	17.7 ms	2.83 M	21.6 ms	4.30 M
Hists	No Dist.	9.51 ms	1.34 M	17.7 ms	4.98 M	21.6 ms	7.66 M
Hist	With Dist.	1.66 ms	772 k	5.71 ms	3.08 M	8.45 ms	4.80 M
Hist	No Dist.	1.76 ms	1.31 M	5.82 ms	5.23 M	8.56 ms	8.16 M
Pattern	All Lines	26.4 ms	19.1 M	152 ms	76.5 M	197 ms	120 M
Pattern	20 Lines	2.80 ms	289 k	5.31 ms	1.16 M	6.45 ms	1.81 M
Pattern	10 Lines	1.62 ms	112 k	3.02 ms	449 k	3.78 ms	702 k

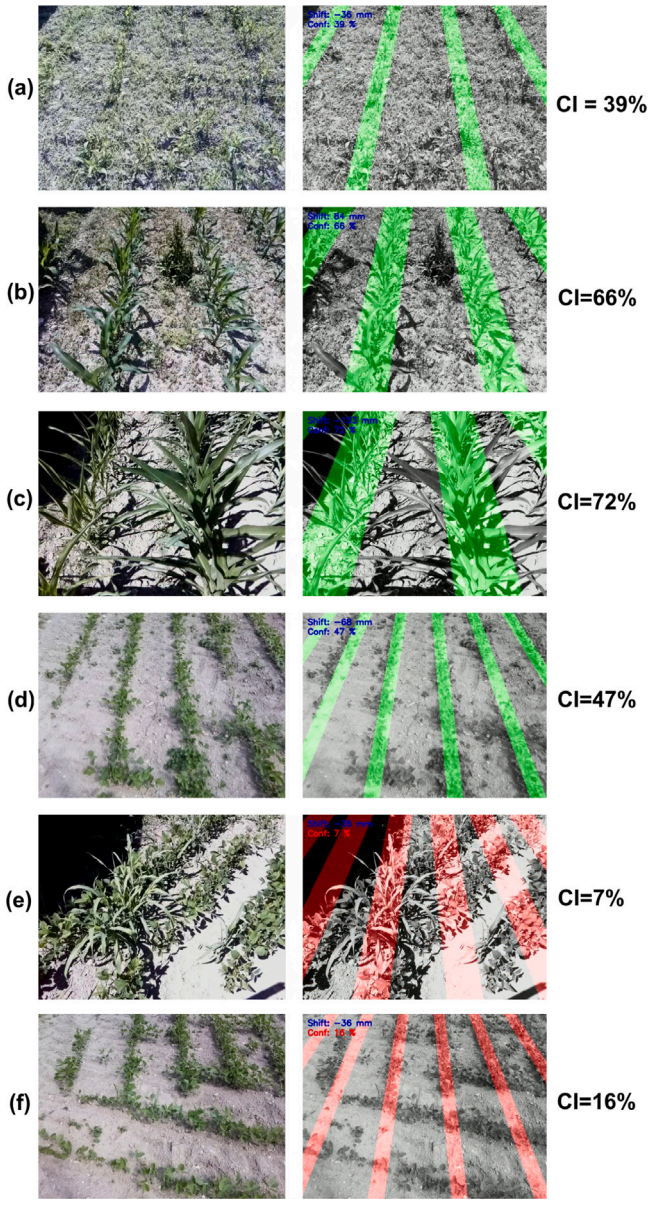

Fig. 31. Pattern matching algorithm applied to real images: matching rows and confidence index.

3.6. Next developments

The proposed results refer to a dataset of static images, on which indices have been evaluated statistically. The position control system will be an electromechanical real-time system operating on the video stream of a camera.

There is an opportunity to also exploit inter-frame or time information.

By averaging a certain number of frames, preferably after having segmented them, it is possible to obtain a blurred image in which the persistent structures are highlighted: in this way it is possible to emphasize the crop rows to the detriment of the weeds, which typically assume irregular positions. However, the time length of the average poses two limits: it introduces a delay in the measurement and must be less than the time of the fastest movement to be detected.

The deviation measurements and confidence index are processed at regular time intervals, following the frame rate of the camera. It may be necessary to filter these values in order to eliminate out-of-range measurements, which can be achieved with a low-coefficient IIR filter.

The analysis of the benefits of these proposals is postponed to a subsequent work in which the dynamic behavior of the proposed algorithm is evaluated.

The presented study focused on the image measurement system. An integral part of the scope is the electromechanical control system that will be used, which introduces delays and vibrations related to the mechanical characteristics of the drive.

Both to validate the overall system and to measure the improvement achievable by varying the parameters of the algorithm or by exploring the previous proposal, it is advisable to design an automatic system for collecting field measurements.

A basic solution involves measuring the movement of the cultivator frame to compare it with the measurements of the vision system, highlighting delays and errors in the control ring. This solution, however, suffers from not having a common absolute reference; in fact two relative measurements are compared, one referring to the theoretical center row for the vision system, and one referring to the frame for the cultivator, which is affected by the tractor-field position.

For a complete analysis it is advisable to find a system that measures the position of the cultivator in relation to the plants, but with a technology different from the vision system, for example using instrumentation based on ToF or Lidar sensors, choosing systems with measurement error well below the image method under consideration.

4. Conclusion

The proposed work considers some vision algorithms for an optical guide in crops rows, analyzing in terms of accuracy some of their variants and simplifications.

A first common aspect regards the detection of plants in the image it is highlighted, experimentally, the opportunity to threshold the image in the Lab color space in order to identify a tint of interest: selectively for the green-red color ratio, which identifies the cultivation to be detected, and wider for the blue–yellow color ratio, which takes into account the different color temperatures in the image.

With regard to the vision algorithms, it should be noted that the proposal to correct the image prospectively and analyze a single vertical histogram instead of multiple bands improves the measurement resolution, especially when the seeding pattern becomes disturbed by infestations or missing plants, together with a speed up of the code. The results also show that working on an image affected by optical

distortion introduces a smaller measurement error by analyzing a single histogram with adjusted perspective since the distortions at the edges are partly eliminated from the histogram, precisely by the perspective correction, and partly mediated; conversely, analyzing the histograms of several bands, the peripheral bands contribute with error to the final displacement estimate. However, both algorithms show an increase in the measurement uncertainty as weeds or missing plants grow, arriving in the simulated cases at uncertainties over 40 mm, values that are too relevant for the weeding application.

The proposed algorithm based on pattern matching, on the other hand, is not affected by optical distortion measurements error, or does not need its correction, since the pattern can be distorted in an appropriate way. It shows greater accuracy than histogram-based methods, with measurement precision in the order of 25–30 mm even in the presence of many weeds or missing plants; also r^2 index shows greater measurement accuracy. On the tested processor system it is slower than the histogram methods but it has been shown, even without apply to dedicated hardware for calculating the correlation, that is possible to reduce the computational load by analyzing only a few lines of the image while maintaining an adequate measurement resolution, in any case within 30 mm for simulated normal use cases.

Funding

This research was funded by the grant POR-FESR 2014–2020, Activity 1.3.a "New intelligent machines and systems for treatment on foliar apparatus, irrigation and sowing". The specific activity by Francesco Marinello is supported by the Agritech National Research Centre funded from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR)—MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4—D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

CRediT authorship contribution statement

Luca De Bortoli: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Stefano Marsi:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Formal analysis, Data curation, Conceptualization. **Francesco Marinello:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Funding acquisition, Data curation, Conceptualization. **Paolo Gallina:** Visualization, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

The Laboratory for Advanced Mechatronics—LAMA FVG is gratefully acknowledged for technical support.

References

Aden Darge, D.Z..P.Y.K.C., 2019. Multi color image segmentation using L*A*B* color space. Int. J. Adv. Eng. Manag. Sci. 5, http://dx.doi.org/10.22161/ijaems.5.5.8.

Ban, C., Wang, L., Chi, R., Su, T., Ma, Y., 2024. A camera-LiDAR-IMU fusion method for real-time extraction of navigation line between maize field rows. Comput. Electron. Agric. 223, 109114. http://dx.doi.org/10.1016/j.compag.2024.109114, URL https://www.sciencedirect.com/science/article/pii/S0168169924005052.

Burgos-Artizzu, X.P., Ribeiro, A., Guijarro, M., Pajares, G., 2011. Real-time image processing for crop/weed discrimination in maize fields. Comput. Electron. Agric. 75 (2), 337–346. http://dx.doi.org/10.1016/j.compag.2010.12.011, URL https://www.sciencedirect.com/science/article/pii/S0168169910002620.

Correa, C., Valero, C., Barreiro, P., 2011. Row crop's identification through hough transform using images segmented by robust fuzzy possibilistic C-means. Inteligencia Artif. Rev. Iberoamericana De Inteligencia Artif..

de Silva, R., Cielniak, G., Gao, J., 2024. Vision based crop row navigation under varying field conditions in arable fields. Comput. Electron. Agric. 217, 108581. http://dx.doi.org/10.1016/j.compag.2023.108581, URL https://www.sciencedirect.com/science/article/pii/S0168169923009699.

Devore, J.L., 2011. Probability and Statistics for Engineering and the Sciences, eighth ed. Cengage Learning, Boston, MA, pp. 508–510.

Duarte-Correa, D., Rodríguez-Reséndiz, J., Díaz-Flórez, G., Olvera-Olvera, C.A., Álvarez-Alvarado, J.M., 2023. Identifying growth patterns in arid-zone onion crops (allium cepa) using digital image processing. Technologies 11 (3), http://dx.doi.org/10.3390/technologies11030067, URL https://www.mdpi.com/2227-7080/11/3/67.

Egli, D., 2023. Yield improvement and yield components: A comparison of corn and soybean. Crop Sci. http://dx.doi.org/10.1017/wsc.2023.50.

Ericson, S., Åstrand, B., 2010. Row-detection on an agricultural field using omnidirectional camera. In: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 4982–4987. http://dx.doi.org/10.1109/IROS.2010.5650964.

Fisher, R., Perkins, S., Walker, A., Wolfart, E., 2003. Hough transform. http://homepages.inf.ed.ac.uk/rbf/HIPR2/hough.htm.

García-Santillán, I., Peluffo, D., Caranqui, V., Pusdá, M., Garrido, F., Granda-Gudiño, P., 2018. Computer vision-based method for automatic detection of crop rows in potato fields. ISBN: 978-3-319-73449-1, pp. 355–366. http://dx.doi.org/10.1007/978-3-319-73450-7_34.

Hague, T., Tillett, N., 2001. A bandpass filter approach to crop row location and tracking. Mechatronics 11, 1–12. http://dx.doi.org/10.1016/S0957-4158(00)00003-9.

Ji, R., Qi, L., 2011. Crop-row detection algorithm based on random hough transformation, 54 (3–4), http://dx.doi.org/10.1016/j.mcm.2010.11.030.

Jiang, G., Wang, Z., Liu, H., 2015. Automatic detection of crop rows based on multi-ROIs. Expert Syst. Appl. 42 (5), 2429–2441. http://dx.doi.org/10.1016/j.eswa.2014.10.033, URL https://www.sciencedirect.com/science/article/pii/S0957417414006575.

Li, Y., Huang, Z., Cao, Z., Lu, H., Wang, H., Zhang, S., 2020. Performance evaluation of crop segmentation algorithms. IEEE Access 8, 36210–36225. http://dx.doi.org/10.1109/ACCESS.2020.2969451.

Li, C., Lee, C., 1993. Minimum cross entropy thresholding. Pattern Recognit. 26 (4), 617–625. http://dx.doi.org/10.1016/0031-3203(93)90115-D.

Li, Z., Wang, S., Sun, J., 2009. Image segmentation in object recognition of mature eggplant. 40, 105–108+96.

Liu, T., Zheng, Y., Lai, J., Cheng, Y., Chen, S., Mai, B., Liu, Y., Li, J., Xue, Z., 2024. Extracting visual navigation line between pineapple field rows based on an enhanced YOLOv5. Comput. Electron. Agric. 217, http://dx.doi.org/10.1016/j.compag.2023.108574.

Machleb, J., Peteinatos, G., Kollenda, B., Andújar, D., R., G., 2020. Sensor-based mechanical weed control: Present state and prospects. Comput. Electron. Agric. 176, http://dx.doi.org/10.1016/j.compag.2020.105638.

Montalvo, M., Pajares, G., Guerrero, J., Romeo, J., Guijarro, M., Ribeiro, A., Ruz, J., Cruz, J., 2012. Automatic detection of crop rows in maize fields with high weeds pressure. Expert Syst. Appl. 39 (15), 11889–11897. http://dx.doi.org/10.1016/j.eswa.2012.02.117, URL https://www.sciencedirect.com/science/article/pii/S0957417412003806.

Nicholls, P., Green, R., 2017. Detecting tramways in crops for robot navigation. https://github.com/petern3/crop_row_detection/blob/master/Detecting%20Tramways%20in%20Crops%20for%20Robot%20Navigation.pdf (Accessed April 2023).

Okamoto, H., Hata, S., Kataoka, T., Terawaki, M., 2001. Automatic weeding cultivator using crop-row detector. IFAC Proc. Vol. 34 (19), 117–122.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. 9 (1), 62–66. http://dx.doi.org/10.1109/TSMC.1979.4310076.

Roberts, Roberts algorithm.

Romeo, J., Pajares, G., Montalvo Martínez, M., Guerrero, J., Guijarro, M., Ribeiro, A., 2012. Crop row detection in maize fields inspired on the human visual perception. Sci. World J. 2012, 484390. http://dx.doi.org/10.1100/2012/484390.

Singh, M., Thapa, R., Singh, N., Mirsky, S., Acharya, B., Jhala, A., 2020. Does narrow row spacing suppress weeds and increase yields in corn and soybean? A meta-analysis. Weed Sci. 176, http://dx.doi.org/10.1016/j.compag.2020.105638.

Søgaard, H., Olsen, H., 2003. Determination of crop rows by image analysis without segmentation. Comput. Electron. Agric. 38 (2), 141–158. http://dx.doi.org/10.1016/S0168-1699(02)00140-0, URL https://www.sciencedirect.com/science/article/pii/S0168169902001400.

Tillett, N., Hague, T., Miles, S., 2002. Inter-row vision guidance for mechanical weed control in sugar beet. Comput. Electron. Agric. 33 (3), 163–177. http://dx.doi.org/10.1016/S0168-1699(02)00005-4.

Vidović, I., Cupec, R., Hocenski, v., 2016. Crop row detection by global energy minimization. Pattern Recognit. 55 (C), 68–86. http://dx.doi.org/10.1016/j.patcog.2016.01.013.

Wang, T., Chen, B., Zhang, Z., Li, H., Zhang, M., 2022. Applications of machine vision in agricultural robot navigation: A review. Comput. Electron. Agric. 198, http://dx.doi.org/10.1016/j.compag.2022.107085.

Winterhalter, W., Fleckenstein, F.V., Dornhege, C., Burgard, W., 2018. Crop row detection on tiny plants with the pattern hough transform. IEEE Robot. Autom. Lett. 3 (4), 3394–3401. http://dx.doi.org/10.1109/LRA.2018.2852841.

Woebbecke, D.M., Meyer, G.E., Bargen, K.V., Mortensen, D.A., 1994. Color indices for weed identification under various soil, residue, and lighting conditions. Trans. ASABE 38, 259–269.

Zhang, L., Zhang, R., Li, L., Ding, C., Zhang, D., Chen, L., 2022. Research on virtual Ackerman steering model based navigation system for tracked vehicles. Comput. Electron. Agric. 192, 106615. http://dx.doi.org/10.1016/j.compag.2021.106615, URL https://www.sciencedirect.com/science/article/pii/S0168169921006323.

Zheng, Z., Hu, Y., Li, X., Huang, Y., 2023. Autonomous navigation method of jujube catch-and-shake harvesting robot based on convolutional neural networks. Comput. Electron. Agric. 215, http://dx.doi.org/10.1016/j.compag.2023.108469.

Zong, Z., Liu, G., Zhao, S., 2020. Real-time localization approach for maize cores at seedling stage based on machine vision. Agronomy 10 (4), http://dx.doi.org/10.3390/agronomy10040470, URL https://www.mdpi.com/2073-4395/10/4/470.