

# TOMATO DEVELOPMENT MONITORING IN AN OPEN FIELD, USING A TWO-CAMERA ACQUISITION SYSTEM

Ujjwal Verma<sup>1,2</sup>, Florence Rossant<sup>1</sup>, Isabelle Bloch<sup>2</sup>, Julien Orensanz<sup>3</sup>,  
Denis Boisgontier<sup>3</sup> and Marthe Lagarrigue<sup>1</sup>

<sup>1</sup>ISEP, Paris

<sup>2</sup>Institut Mines-Telecom, Telecom ParisTech, CNRS LTCI, Paris

<sup>3</sup>Cap2020, Gironville sur Essonne, France

## ABSTRACT

Monitoring the growth of crops and estimating their yield allows the farmer to better manage the resources after the harvest. The existing methods for monitoring the growth of crops depend on remote sensing data or soil and weather information, which are prone to error due to adverse climatic conditions or insufficient information.

Here we present an alternative approach, where the growth of the crop (tomato) is monitored from images captured in an open field. For this, the tomatoes are identified in the image using a segmentation procedure and their size is measured. This is a challenging task because of severe occlusions and poor contrast in the images. In order to increase the robustness of the segmentation procedure (based on active contours) and simplify the size estimation, we approximate the tomatoes as spheres in the 3D space, hence as ellipses in the image space. This model enables us to integrate *a priori* information about the shape of the object to be segmented and to avoid a complete reconstruction of the 3D scene. The automatic segmentation was evaluated by comparing the results with manual segmentation. For the cases with a reasonable amount of occlusion (less than 30%), good results were obtained, with an average relative error of 6.46% (expressed as a percentage of the tomato size). The metric reconstruction has also been evaluated. It was observed that the error on the estimated tomato radius was less than 5% for 91% of the cases. Finally, the complete system was tested. The size of the tomatoes was correctly estimated in 80% of the cases. In addition, preliminary studies have showed that the actual volume of tomatoes can be estimated from the calculated sphere radius, using a correction factor.

**Keywords:** Precision agriculture, Image segmentation, Metric reconstruction

## INTRODUCTION

Monitoring the growth of crops and estimating the yield and the date of harvest are valuable information when farming open field tomatoes. It allows better management of resources (such as storage requirements, transportation) after the harvest. It also enables the farmer to better negotiate terms and conditions for crop insurance. Monitoring growth during tomato's early stages is also interesting to assess plant stress or abnormal development.

One of the existing methods to estimate the yield uses remote sensing data (vegetation indices). However, the quality of the acquired data may decrease due to adverse climatic conditions (clouds, etc.) [Prasad et al., 2006, Mkhabela et al, 2011]. Growth modeling is another method that uses and combines information about the cultivated crop variety, soil and weather, in order to model the crop growth and compute a theoretical estimate of the yield [Zhao and Pei, 2013]. However, this method considers an ideal case with no infected plants. Finally, very few works address the problem of crop monitoring from images. For example, Aggelopoulou and Stajanko process images of apple orchards [Aggelopoulou et al, 2011, Stajanko and Cmelik, 2005]. However, these approaches are limited to a controlled environment where complex scenarios such as occlusion are not considered. For example, in one of the proposed methods [Aggelopoulou et al, 2011], the observed scene was modified (using a black cloth placed behind the tree) in order to simplify the image processing task.

In this work, we present a different approach to monitor the growth of tomatoes using images acquired in an open field during the entire agriculture season while keeping the cost as low as possible. For this, two cameras (left and right) were installed and images were acquired at regular intervals (Figure 1). The entire image processing system is made up of two main parts: first we identify the tomatoes using a segmentation procedure, and then we estimate their size, from both right and left images, by applying a metric reconstruction scheme. The yield of the field can be estimated from this information.



**(a) Left image**

**(b) Right image**

**Figure 1. Images of the tomatoes cultivated in an open field.**

Note that this work is a part of a larger collaborative project, MCUBE, where the multimedia processing capabilities are integrated into a classical machine to machine (M2M) system, allowing the user to remotely monitor the field.

## CHALLENGES

One of the challenges of the system results from occlusions: most of the tomatoes are partially hidden by other tomatoes and/or leaves (Figure 2). Moreover color information is not of much use, because all tomatoes have the same color as the leaves during most of agricultural season time, except at the end of the maturation. In some images a shadow, created by the leaves or the tomato itself, can be observed. As a result, a portion of the contour is not apparent (Figure 2). This results in an ambiguity on the actual position of the contour. The second main challenge concerns the metric reconstruction from the pairs of images captured by the two cameras. This task requires extracting pairs of points, one in each image, that correspond to the same point in the 3D space. This matching process is difficult, given the complexity of the scene.

## PROPOSED SYSTEM

In order to overcome the difficulties mentioned above, we propose to introduce *a priori* shape and temporal information throughout our system.

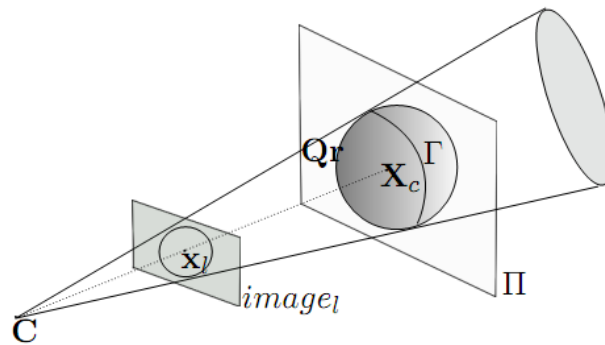
**Shape information:** We propose first to assume that a tomato can be approximated by a sphere in the 3D space. Using the properties of projective geometry, one can easily demonstrate that the image of a sphere is an ellipse [Hartley and Zisserman, 2004].

Figure 3 shows the contour generator  $\Gamma$  of a sphere  $Qr$  of center  $X_C$ : this is the circle of same center and same radius that lies in the plane  $\Pi$  orthogonal to the line joining  $X_C$  and the camera center  $C$ . This circle projects into an ellipse of center  $x_l$  in the image plane. This *a priori* shape information is incorporated in the segmentation procedure thus increasing its robustness with respect to occlusion. Moreover, the spherical hypothesis enables us to simplify dramatically the procedure for size estimation, which can be deduced from the estimated radius of the sphere. This eliminates the need for a full 3D reconstruction of the scene.

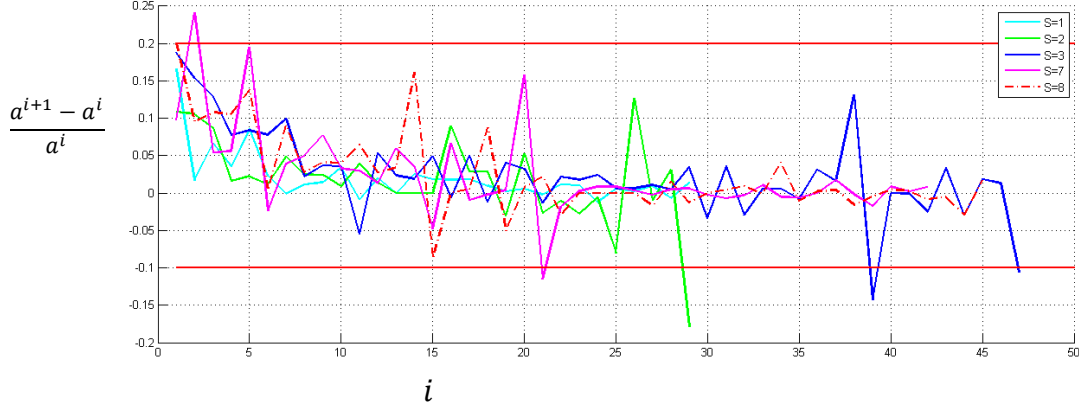


**Figure 2. Occlusions due to leaves and branches (left and center), shadowing effect observed in the lower part of the contour (right).**

**Temporal information:** There is little tomato growth during a given day. Thus only two images per day are studied (one for each camera taken at the same time in the day) creating a sequence of images for a particular tomato. We performed the manual segmentation of five tomatoes (elliptic approximation) and studied the evolution of the ellipse parameters along the time axis. For instance, Figure 4 shows the evolution of the major axis length of the five tomatoes during the entire agriculture season. It can be observed that there is little growth between two consecutive days. Moreover, it has also been observed that there is little movement of the tomato between two consecutive days. However, this movement is random and very difficult to predict especially in case of strong wind and rain. This temporal information is also incorporated in the segmentation step by considering those ellipses whose parameters lie within an acceptable range knowing the segmentation validated in the previous image.



**Figure 3. Image of a sphere.**



**Figure 4. Evolution of the length of the major axis of manually segmented ellipses obtained on five selected sequences ( $S = 1, 2, 3, 7, 8$ ).**

**Active contour model:** The segmentation procedure is based on an active contour model [Kass et al, 1988, Xu et al, 2000] with shape constraint. This model deforms the contour iteratively from its initial position towards the edges of the object by minimizing an energy functional.

The total energy functional associated with a contour  $v$  is usually modeled as the sum of three energy terms: the internal energy  $E_{int}$ , the image energy  $E_{Im}$  and the external energy  $E_{ext}$ :

$$E_{Total}(v) = E_{int}(v) + E_{Im}(v) + E_{ext}(v)$$

The internal energy term controls the physical properties (stiffness, elasticity) of the evolving contour. The image energy is derived from image data and drives the contour towards strong gradients in the image. *A priori* knowledge about the object to be segmented is included in the external energy term. In our work, we propose to constrain the evolution of the contour so that it remains close to a reference ellipse.

Our model relies on a parametric representation of the curves, since ellipses are easily represented in this way. Let us denote by  $z_e(\theta) = r_e(\theta)e^{j\theta}$ ,  $\theta \in [0, 2\pi]$ , the reference ellipse represented in polar coordinates, with origin at its center. A point  $z(\theta)$  lying on the active contour can be defined in the same way as

$$z(\theta) = r(\theta)e^{j\theta}$$

We propose to define the total energy with elliptic shape prior as:

$$E_T(r, r_e) = \int_0^{2\pi} \frac{\alpha}{2} |r'(\theta)|^2 d\theta + \int_0^{2\pi} E_{Im}(r(\theta)e^{j\theta}) d\theta + \frac{\psi}{2} \int_0^{2\pi} |r(\theta) - r_e(\theta)|^2 d\theta$$

In this equation, the first term represents the internal energy while the second term represents the image energy, calculated from the diffusion of the gradient vectors [Xu and Prince, 1998]. The last term ( $\frac{\psi}{2} \int_0^{2\pi} |r(\theta) - r_e(\theta)|^2 d\theta$ ) can be viewed as the mean square error between the evolving contour and the reference ellipse.

Thus, its minimization ensures that the active contour remains close to the reference ellipse.

The minimization of the energy functional is performed through an iterative process, the reference shape  $z_e(\theta)$  being regularly updated from points extracted from the evolving contour  $z(\theta)$ .

## SEGMENTATION ALGORITHM

**Method:** We consider here a sequence of images acquired from the left camera; the same processing is applied for images given by the right camera. Given the challenges of the system, we assume that the segmentation in image  $i$  has been validated by an operator, before segmenting the tomatoes in the  $(i + 1)^{th}$  image. Indeed, this validation task is not heavy and enables us to rely on the ellipse parameters obtained in the previous image to find the new ones, through regularization in time and space. The segmentation procedure is described in detail in [Verma et al, 2014]. A brief overview of the method is presented below.

Since tomatoes move between two consecutive images, we first start by updating the position of the tomato in the  $(i + 1)^{th}$  image. The segmentation procedure is based on the active contour model presented above, which requires to be initialized. This initial contour is computed in two steps: first several candidate ellipses are determined from gradient information through the RANSAC algorithm. Then additional region information is extracted in order to select one ellipse among these estimates as the initial contour. The active contour with shape constraint is then applied. Finally, four different ellipse estimates are computed ( $Ell_{f1}, Ell_{f2}, Ell_{f3}, Ell_{f4}$ ), based on point selection rules that aim at extracting the points of the final curve that actually lie on non-occluded parts of the tomato boundary. The operator only has to select one estimate as the final segmentation. Note that we also compute the region of potential occlusion using region information, and modified the active contour model so that the effect of occlusion is minimized during the evolution of the active contour.

Although the image processing algorithms used in our segmentation procedure are classical ones, the main originality of the approach results from the combination of three types of information: gradient information, region information and prior information regarding the shape and size evolution of tomatoes.

**Results:** The cameras were installed in open fields of tomatoes, capturing a delimited region (Figure 1). The same setup was used for three agricultural seasons (April-August, 2011, 2012 and 2013). The vegetation was very different in the three cases, although the variety of the cultivated tomatoes was the same, because of different climatic conditions. We studied 21 tomatoes, covering different sites and different seasons, thus ensuring variability.

In order to perform a quantitative evaluation of the proposed approach, we compared the segmentations obtained with our algorithm with manual

segmentations (approximated by ellipses). Let us assume that the manual segmentation  $M$  is an ellipse representing the tomato, and that the automatic segmentation is represented by a set of equally-spaced points  $A = \{(x_a, y_a), a = 1, 2, \dots, a_n\}$ . The mean distance  $D_{mean}(A, M)$  between the two contours  $A$  and  $M$  is defined as:

$$D_{mean}(A, M) = \frac{1}{a_n} \sum_{a=1}^{a_n} d_e((x_a, y_a), M)$$

where  $d_e((x_a, y_a), M)$  is the Euclidean distance between  $(x_a, y_a)$  and the ellipse  $M$ . In order to better interpret the results, we express the measured error as a percentage of the tomato size. Let us denote by  $a^i, b^i$  the lengths of the semi-major and semi-minor axes in the  $i^{th}$  image (given by the manual segmentation). The normalized mean distance error is defined by:

$$D_{meanR}^i = \frac{D_{mean}^i}{r^i} 100 \quad \text{with } r^i = \frac{a^i + b^i}{2}$$

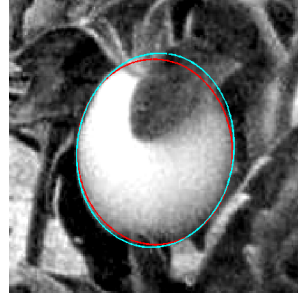
An evaluation of the segmentation procedure on the entire image data set would be meaningless, given the heterogeneity of the image quality especially in terms of occlusion degree. Therefore, the influence of the amount of occlusion on the estimated radius was studied beforehand, considering the complete system applied on the ideal case of a spherical object. Using this information, the image dataset was divided into three categories, depending on the amount of occlusion. An image belongs to category 1 if the percentage of occlusion is less than 30%, to category 2 if the percentage of occlusion is between 30% and 50%, and to category 3 if the percentage of occlusion is greater than 50%.

Table 1 shows the mean and standard deviation of  $D_{meanR}$  for the 21 tomatoes that have been followed up. In this table, only the images of category 1 have been considered and results are given for the fourth approximation output by the algorithm, denoted by  $Ell_{f4}$ . It is worth noting the low mean and lower standard deviation in most cases, which demonstrates both the accuracy and the robustness of our algorithm (Figure 5).

However, a slightly higher distance measure was observed in some images. The imprecision of the obtained contour is mainly due to blurring or shadows, that result in a strong ambiguity on the actual position of the tomato boundary (Figure 6). Moreover, for the images acquired during the agricultural season 2013 (sequences  $S = 12, \dots, 21$ ), the sizes of tomatoes were comparatively smaller than that of agricultural seasons 2011 and 2012. Since the error measures  $D_{meanR}$  are normalized by the size of the tomato ( $r^i$ ), higher values are generally observed for the images of these sequences. For example, Figure 7 shows the obtained segmentation  $Ell_{f4}$  on the 3<sup>rd</sup> image of sequence  $S = 17$ . The distance measure normalized by the size of the tomato is  $D_{meanR} = 9.89\%$ . However, the distance measure expressed in pixels is not so important ( $D_{mean} = 2$  pixels).

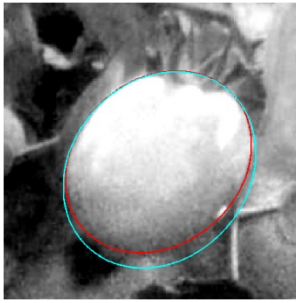


(a) ( $D_{meanR} = 0.94\%$ )

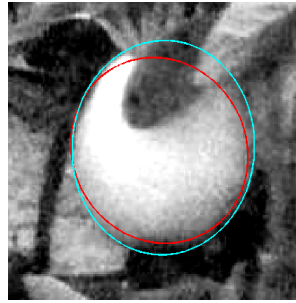


(b) ( $D_{meanR} = 3.05\%$ )

**Figure 5. Automatic segmentation ( $Ell_{f4}$ ) (red) along with manual segmentation (cyan) for two images of category 1.**

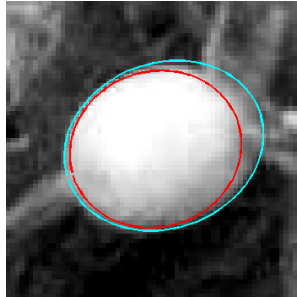


(a) ( $D_{meanR} = 3.38\%$ )



(b) ( $D_{meanR} = 7.94\%$ )

**Figure 6. Automatic segmentation ( $Ell_{f4}$ ) (red) along with manual segmentation (cyan) for two images of category 1 with blurred contours.**



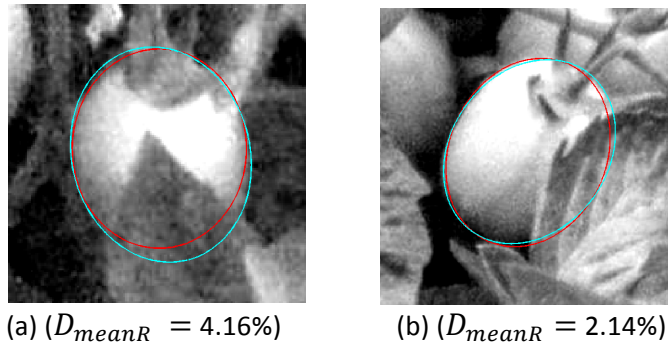
**Figure 7. Automatic segmentation (red) on the 3<sup>rd</sup> image of sequence  $S = 17$ . The distance measure is  $D_{meanR} = 9.89\%$ .**

For the images of category 2, which contain a significant amount of occlusion (between 30% and 50%), good results were obtained for most sequences, with an average distance error less than 10% for 73% of the images. Figure 8 shows some examples where a good segmentation was obtained on these images despite the presence of occlusion by leaves.

Until now we have only studied one of the ellipses ( $Ell_{f4}$ ) among the four ellipse estimates produced by our algorithm. This ellipse estimate is not necessarily the best ellipse among the four estimates for all images and was only



selected for illustrative purpose. In general, for images of good quality with a low amount of occlusion, the four ellipses are very similar. In less ideal conditions, one of the four ellipses generally better matches the actual tomato boundary than the others. Let us denote by  $Ell_{opt}$  the ellipse that has been selected by the operator among the four ellipse estimates, as expected in normal conditions of use. The distance measures for this ellipse are shown in Table 1. It can be observed that these measures are significantly lower as compared to those for  $Ell_{f4}$ , with a difference ranging from 0.01% to 1.88%. This shows that there is always an ellipse, among the four estimates, that best represents the tomato and the operator only has to select this ellipse as the final segmentation.



**Figure 8. Automatic segmentation  $Ell_{f4}$  (red) along with manual segmentation (cyan) for two images of category 2.**

**Table 1: Mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of  $D_{meanR}$  measured on the 21 sequences, for images of category one. One of the four ellipses produced by the segmentation algorithm is considered ( $Ell_{f4}$ ) along with the best ellipse selected by the operator ( $Ell_{opt}$ ).**

	Number of images in category 1	$Ell_{f4}$		$Ell_{opt}$	Difference in $\mu_{D_{meanR}} (Ell_{f4} - Ell_{opt})$
		$\mu_{D_{meanR}}$	$\sigma_{D_{meanR}}$	$\mu_{D_{meanR}}$	
Sequence 1	26	1.72	0.77	1.37	0.35
Sequence 2	4	1.85	0.46	1.66	0.19
Sequence 3	21	3.4	2.24	2.89	0.51
Sequence 4	14	2.73	1.92	2.22	0.51
Sequence 5	5	4.81	1.3	4.82	-0.01
Sequence 6	0	-	-	-	-
Sequence 7	25	1.88	0.65	1.72	0.16
Sequence 8	20	6.07	5.75	5.46	0.61
Sequence 9	1	5.26	0	5.39	-0.13
Sequence 10	5	2.25	0.56	1.86	0.39
Sequence 11	4	11.81	4.99	10.54	1.27
Sequence 12	19	4.74	1.33	4.25	0.49
Sequence 13	5	41.5	16.55	40.48	1.02
Sequence 14	4	9.57	2.35	9.18	0.39
Sequence 15	0	-	-	-	-
Sequence 16	21	4.68	1.08	4.46	0.22
Sequence 17	20	11.78	2.44	11.56	0.22
Sequence 18	23	14.18	20.06	13.94	0.24
Sequence 19	0	-	-	-	-
Sequence 20	5	8.76	5.38	6.88	1.88
Sequence 21	25	7.34	3.18	7.12	0.22

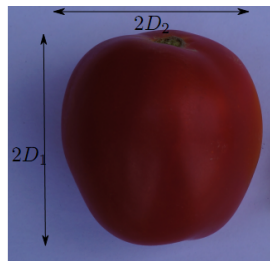
## SIZE ESTIMATION

Once the tomatoes have been identified in both images, we then estimate their size. This requires the knowledge of the camera parameters which are calculated once for all at the beginning of the season, using a calibration pattern [Bouguet, 2013]. Knowing the elliptic approximations of the tomato in the left and right images, a triangulation procedure recovers the center of the sphere in the 3D space [Bouguet, 2013]. Each image is then considered independently of the other one in order to recover the 3D points of the contour generator and then estimate the radius of the sphere. Geometrical properties are used to this aim. Indeed, the coordinates of the 3D points can be recovered by calculating the intersection between the plane  $\Pi$  (Figure 3) and the rays back-projected from the ellipse points of the image plane. This procedure provides two different estimates of the radius, one for each image. Finally a joint optimization procedure enables us to estimate the sphere radius.

**Results:** In order to evaluate our size estimation procedure, images of ten tomatoes ( $T = 1, 2, \dots, 10$ ) were acquired in the laboratory. Two different distances with respect to the camera (Positions A and B, corresponding to the first two rows of tomatoes cultivated in an open field) and two different heights ( $H = 10$  cm and 30 cm) with respect to the ground were considered. Also, for each position and each given height, the tomatoes were observed at three different orientations, which were not necessarily identical for all positions and heights. The acquisition system is identical to the one used in the open field. The images were manually segmented as we wished to evaluate the metric reconstruction only.

In our calculation, it is assumed that the tomato is a sphere in the 3D space. However, in reality, the tomato is not a perfect sphere. Figure 9 shows the two reference values  $D_1$  and  $D_2$  (ground truth) that were measured. These reference values were compared with the radius  $r_{est}$  estimated using the proposed method. Let us define two different error percentages with respect to the different reference values as:

$$PE_{D_i} = \frac{|D_i - r_{est}|}{D_i} 100, \quad i = 1, 2$$



**Figure 9. Two reference values**

Table 2 shows the  $PE_{D_1}$  percentages that have been obtained for all positions and orientations. This error is always less than 10% and most (91%) of the values are less than 5%, which demonstrates the robustness of our method. The comparison with the smaller reference value  $D_2$  gives slightly higher  $PE_{D_2}$  percentages in most of the cases. This is logical because a sphere of radius equal to  $D_1$  covers the entire tomato. However, depending on the orientation of the tomato in the 3D space one or the other reference value can be best approximated.

**Table 2: Percentage error  $PE_{D_1}$  for the tomatoes  $T = 1, \dots, 10$  computed by comparing the estimated radius with the reference value  $D_1$ .**

		T =	1	2	3	4	5	6	7	8	9	10
Pos = A	H = 10	O = 1	2.11	2.68	4.79	1.79	3.27	1.42	4.70	4.03	4.20	6.63
		O = 2	2.47	7.32	3.81	1.10	0.15	1.09	3.80	3.87	5.07	8.46
		O = 3	4.51	5.65	4.66	0.33	1.98	0.94	2.66	5.02	5.01	5.65
	H = 30	O = 1	1.51	4.63	0.34	2.70	1.83	4.61	4.20	1.61	2.10	1.91
		O = 2	0.58	3.97	5.32	0.95	0.86	0.04	2.19	1.14	0.25	0.30
		O = 3	1.00	4.73	5.13	4.53	2.52	3.94	2.49	0.39	1.92	0.37
Pos = B	H = 10	O = 1	4.61	1.59	0.64	0.97	4.23	2.23	0.47	2.21	0.17	4.76
		O = 2	4.89	0.87	0.03	0.42	2.35	1.74	0.03	1.29	2.62	0.88
		O = 3	4.05	0.55	1.71	1.08	0.45	1.72	0.22	3.29	2.60	3.03
	H = 30	O = 1	0.47	1.75	1.69	0.92	0.30	1.11	3.47	1.60	2.38	3.54
		O = 2	3.10	1.75	1.37	0.66	0.79	1.27	0.38	0.00	2.09	1.93
		O = 3	1.62	1.02	0.65	2.43	0.41	1.17	6.07	4.50	0.39	0.74

The volume of the tomato can be computed from its estimated radius based on the spherical hypothesis. However, since a tomato is not a perfect sphere, a correction factor  $\alpha$  was determined experimentally. Then, the corrected volume was estimated by:

$$V_c = \frac{4}{3} \pi (\alpha r_{est})^3$$

The relative error percentage between the corrected volume and the actual volume  $V_a$  was computed as:

$$PE_V = \frac{|V_a - V_c|}{V_a} 100$$

It was observed that the error percentage is less than 15% in 87% of the cases. These very preliminary results show that it may be possible to correct the measurements made with the spherical hypothesis in order to take into account the specific shape of the tomato variety that is cultivated in the field.

## RESULTS: ENTIRE SYSTEM

The size of the tomatoes cultivated in open fields was also estimated for the agricultural season 2013, considering the entire system. Table 3 shows the estimated radius computed using the automatic segmentation selected by the operator ( $Ell_{opt}$ ) along with the reference values and the error percentages. Again, it can be observed that the estimated radius is generally closer to the larger reference value  $D_1$  than to the smaller  $D_2$ , but not always, depending on the orientation of the tomato.

**Table 3: Estimated radius  $r_{est}$  obtained with the proposed method compared with the reference values  $D_1$  and  $D_2$ .**

	$D_1$ (cm)	$D_2$ (cm)	$r_{est}$ (cm)	$PE_{D_1}$	$PE_{D_2}$
Sequence 12	2.53	2.03	2.69	<b>6.17</b>	32.32
Sequence 13	2.07	1.92	2.16	<b>4.61</b>	12.80
Sequence 14	2.09	1.75	1.77	14.89	<b>1.39</b>
Sequence 15	2.51	2.2	2.17	13.57	<b>0.97</b>
Sequence 16	2.48	2.14	2.29	7.52	<b>6.94</b>
Sequence 17	1.44	1.36	1.17	18.65	<b>13.87</b>
Sequence 18	1.89	1.51	2.00	<b>5.62</b>	32.20
Sequence 19	2.57	2.14	2.84	<b>10.46</b>	33.22
Sequence 20	2.13	1.86	2.71	<b>27.50</b>	45.67
Sequence 21	2	1.77	2.09	<b>4.56</b>	18.14

The higher errors are mainly due to imperfect segmentation because of the ambiguity on the actual position of the tomato contour (shadow, blurring, etc.). Note that the proposed method is divided into several steps, and the accuracy of the final estimate depends on the accuracy of each of these steps. For instance, an imprecision in the segmentation procedure and/or on camera parameters could affect the final radius estimate. We studied the error on the radius estimation due to imprecision in the segmentation results. For example, it was observed that for an error of 1 pixel in the length of the major axis, the radius varies between 0.015 cm and 0.09 cm (0.5% to 3%) from its reference value for the different positions in the scene.

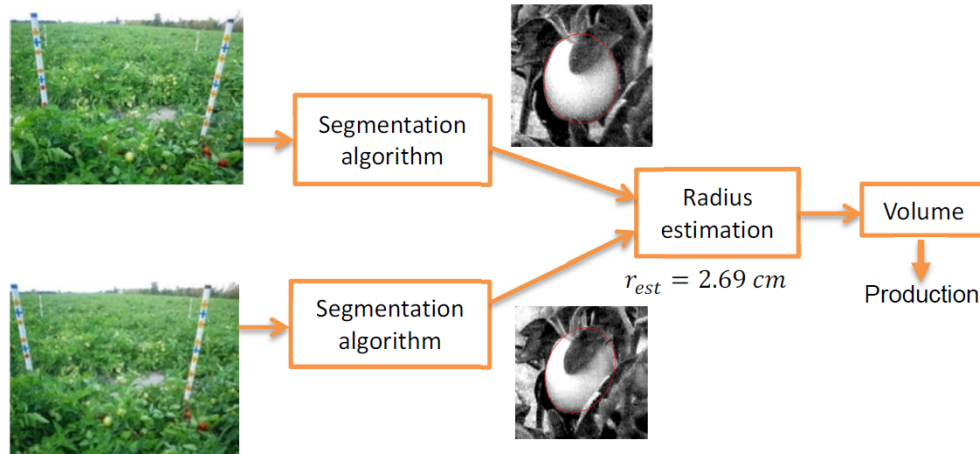
## CONCLUSION

We presented a method for monitoring the growth of tomatoes from images of the field, captured by a set of two cameras. For this, we adapted existing methods to design a complete end to end system according to the requirements of the project. Given the challenges, we proposed to use *a priori* information throughout our system. Figure 10 shows a general overview of the proposed system. First the tomatoes are identified in the images using a segmentation procedure and then their size is calculated. From the estimated size of several tomatoes, the yield of the field can be estimated. Besides, this system allows the user to remotely monitor the growth of the tomatoes throughout the agriculture season. A detailed presentation of the entire system can be found in [Verma, 2014].

The segmentation procedure was evaluated by comparing the results with manual segmentations. The image dataset was divided into three categories according to the amount of occlusion in the images. For the images with occlusion less than 30%, very good results were obtained on most of the images. The average distance between manual and automatic segmentation was less than 10% for 87% of the images. For the images with significant amount of occlusion (between 30 to 50%), this error was less than 10% for 73% of the images.

The metric reconstruction was evaluated as well in laboratory. For that, the estimated tomato size was compared with the actual measured size. We found that

the relative percentage error between the actual size and the estimated size is less than 5% in 91% of the cases.



**Figure 10. Proposed system for yield estimation.**

The proposed system was also used to estimate the size of tomatoes cultivated in an open field during the agriculture season 2013. In this case, the size was correctly estimated in 80% of the cases. Segmentation imprecision is the main cause of error, generally resulting from boundary blurring.

**Perspectives:** Future work will focus on the implementation of the algorithms on a gateway/platform system (Machine to Machine (M2M) architecture), the goal being to get an operational and ergonomic system that enables farmers to monitor easily the tomato growth. The proposed system may also be used to monitor other fruits such as apples.

The present acquisition system can be modified in order to increase the robustness of the system and also gather additional crop information. For instance, the authors in [Sakamoto et al, 2012] have recently shown that the vegetation index computed using visible and near infrared images is closely related to the vegetation index computed using remote sensing data. Therefore, using an additional near infrared camera would allow us to determine crop biophysical parameters. Additionally, it would provide complementary information which can be exploited to improve the robustness of the segmentation procedure. For instance, near infrared images can be used to determine the position of the tomato, thus resulting in a better initialization for the active contours.

## REFERENCES

- Aggelopoulou A., Bochtis D., Fountas S., Swain K., Gemtos T., Nanos G., 2011, Yield prediction in apple orchards based on image processing, *Journal of Precision Agriculture*, 12(3), pp 448-456.
- Bouguet J., Camera calibration toolbox for MATLAB, [http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/), April 2013.
- Hartley R. and Zisserman A., 2004, *Multiple view geometry in computer vision*, Cambridge University Press.
- Kass M., Witkin A., Terzopoulos D., 1988, Snakes: Active contour models, *International Journal of Computer Vision*, 1(4), pp 321-331.
- Mkhabela M., Bullock P., Raj S., Wang S., Yang Y., 2011, Crop yield forecasting on the Canadian Prairies using MODIS NDVI data, *Agriculture and Forest Meteorology*, 151(3), pp 385-393.
- Prasad A., Chai L., Singh R., Kafatos M., 2006, Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation*, 8(1), pp 26-33.
- Stajanko D., and Cmelik Z., 2005, Modelling of apple fruit growth by application of image analysis. *Agriculture Conspectus Scientificus*, 70, pp 59-64.
- Sakamoto T., Gitelson A., Robertson N., Arkebauer T., Wardlow B., Suyker A; Verma S., Shibayama M., 2012, An alternative method using digital cameras for continuous monitoring of crop status, *Agricultural and Forest Meteorology*, 154-155, pp 113-126.
- Verma U., Rossant F., Bloch I., Orensanz J., Boisgontier D., 2014, Shape-based segmentation of tomatoes for agriculture monitoring, *International Conference on Pattern Recognition Applications and Methods (ICPRAM)*, Angers, March 2014, pp 402-411.
- Verma U., 2014, *Analysis of images in precision agriculture: monitoring the growth of tomatoes*, Ph.D. diss., Telecom ParisTech, France.
- Xu C., Pham D., Prince J., 2000, Image segmentation using deformable models *In Handbook of Medical Imaging, Vol 2 Medical Image Processing and Analysis*, pp 175-272.

Xu. C., Prince J.,1998, Snakes, shapes and gradient vector flow, IEEE Transactions on Image Processing, 7(3), pp 359-369.

Zhao H., and Pei Z., 2013, Crop growth monitoring by integration of time series remote sensing imagery and the WOFOST model, International Conference on Agro-Geoinformatics, pp 568-571.