A method of using segmentation of color images and shape factors with frontier rates to identify onion and weeds in field

junwei LU[1], pierre GOUTON[2], jean phillipe GUILLEMIN[3], chenglin MA[4], jean claude COQUILLE[5]

Abstracts - *A new method to solve the problems of overlapping of onion and weed leaves in images* is suggested *in this paper*. *Our method is based on segmentation of color images. It includes three stages : firstly the image is segmented into a lot of small regions. Then these small regions are merged by analyzing the frontier rates and the averages of color indices of pixels belonging to the regions. Thirdly if the small regions can reconstruct a shape of onion leaf or weed leaf, they are merged, otherwise they are not.*

Key Words: color image segmentation, shape analysis , frontier analysis, onion, weeds.

Résumé- Dans cet article nous proposons une méthode pour résoudre le problème du recouvrement des feuilles d'oignons et des mauvaises herbes. Elle comporte trois étapes : Une segmentation des images en beaucoup de petit régions par la couleur; Une première fusion de ces petites régions à partir de l'étude des frontières et selon la couleur; Une seconde fusion suit une analyse de la forme des régions obtenues : si les petites régions peuvent reconstruire une forme de feuille d'oignon ou de feuille de mauvaises herbes, elles sont fusionnées , sinon elles ne le sont pas.

Mots Clés : segmentation d'images couleur , analyse des formes, correspondance de frontières, oignons , mauvaises herbes.

1. Introduction

Decreasing the amount of agriculture chemical is one of the principal aim of Precision Agriculture. By using the analysis of image, one tempts to locate weeds or to measure the proportions of crop and weed leaves in field. In our work, Onions and weeds in field are chosen as objects to be studied. We use the images from some standard photographs. Because the image acquisition is not controlled, difficulties appear when studying plants in field :

- The overlapping of onion and weed leaves.
- Onion and weed and leaves have nearly the same color.
- Color indices of an individual leaf vary in a range greater than the difference between those of two leaves that intersect each other.

Previous studies have concerned color segmentation and shape factors[1] [2]. In this paper we try to solve the specific problems of onion and weed leaves overlapping as it make very difficult distinguish onion and weeds leaves in image.

2. Color indices of plants and soil

By using the statistical analysis, we can obtain averages and standard deviations of the projection of color image in different spaces. Those are:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

----

[1]ENESAD, University of Bourgogne, Jilin University

ENESAD : 26, Bd Dr Petitjean BP 87999, Dijon , France
University of Bourgogne, Dijon, France
Jilin University, Changchun, China
jw.lu@enesad.fr
[2]University of Bourgogne
pierre.gouton@u-bourgogne.fr
[3]ENESAD
jp.guillemin@enesad.fr
[4]Jilin University in China
[5]ENESAD
jc.coquille@enesad.fr

$$\sigma = \left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2\right)^{\frac{1}{2}}$$

μ -- mean value;
x_i -- sampling value;
σ -- standard deviation;
n – sum of sampling points.

It is found that the values of onion and weeds in representation of 2G-R-B[3], hue and intensity layer have a greater difference than in other representations.

3. Analysis of shape factors of onion and weed leaves

3.1 Circularities of onion and weed leaves:

Circularity of onion and weed leaves had been studied by analyzing many images of onion and weed leaves.

$$C(i) = \frac{P(i)^2}{S(i)}$$

where *C(i)* -- circularity of the region *i*
P(i)--perimeter of the region *i*
S(i)-- area. of the region *i*

3.2.Lengths of onion and weed leaves

We define the most long distances between the two points in onion and weed leaves as lengths of onion and weed leaves.

If a(i,j) , b(k,h) are two points in a leaf of onion or weed, then distance of a and b is:

$$Dist(a,b)=\sqrt{(i-k)^2+(j-h)^2}$$

After comparing all distance between the two points in a leaf of onion or weed, the most long distance can be obtained. In fact, one can calculate the distance between the two points in perimeter or in the principal axis of onion and weeds leaves, this can reduce a great deal of quantities of calculation.

3.3 The criterion of recognition

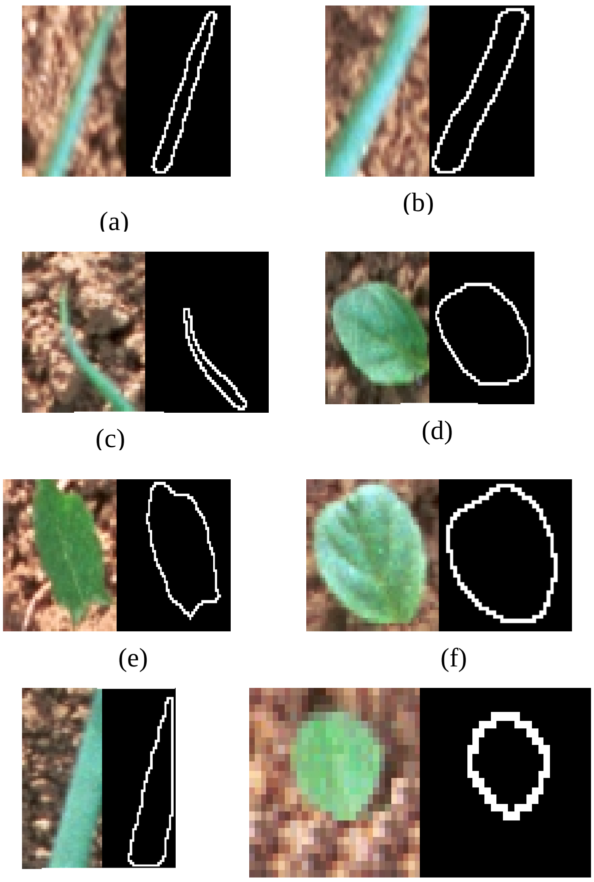
After a great number of weed and onion leaves were analyzed, we give a criterion of recognition of onion and weeds leaves.

that is :

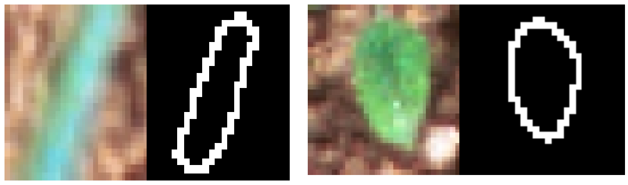
If $L(i)>L_O$
  Then if $C(i) > C_O$ *then i is an onion leaf*
    If $C(i) \le C_O$ *then i is a weed leaf*
If $L(i) \le L_O$
  Then i can not be identified
    ……………………………C.1

i -- object
$L(i)$ *-- length of object i*
$C(i)$ *-- circularity of object i*
L_O , C_O *are thresholds*

Figure 1 shows some examples of onion and weed leaves and their circularities and lengths are listed in table 1 (Weeds: Sharp-leaved fluellen, field bindweed.).



(a) (b) (c) (d) (e) (f) (g) (h)



(i) (j)

FIG. 1: Onion and weeds leaves

TAB. 1: circularities and lengths of leaves in figure 1

Object	Length	Circularity
(a)	62	67
(b)	62	49
(c)	46	74
(d)	43	21
(e)	51	23
(f)	34	19
(g)	74	42
(h)	12	18
(i)	20	27
(j)	20	19

4. Pre-processing color images

By analyzing our images in different color space (RGB, HSI, L*a*b, Ohta), we discover that our image in HSI space have better features than in other space, so HSI space is chosen to study our problems. Comparing onion and weeds characteristics with those of non-plants in different layers (hue, luminosity, saturation), (see table 1), we observe that there are some differences between values obtained from the living plants(onion and weeds) and those from the non-plants (ground, stones, and dead leaves) in layers of hue, saturation and luminosity.

The following algorithm is used to discard all non-plants and to conserve only living plants in the image :

If $\| H(i,j) - \mu_v(H) \| \le K_h * \sigma_v(H)$
  and $\| S(i,j) - \mu_v(S) \| \le K_s * \sigma_v(S)$
  and $\| I(i,j) - \mu_v(I) \| \le K_i * \sigma_v(I)$
  then $H(i,j)$, $S(i,j)$ *and* $I(i,j)$ *are maintained*

otherwise $H(i,j) = 0$ *and* $S(i,j)=0$ *and* $I(i,j)=0$

$H(i,j)$, $S(i,j)$ and $I(i,j)$ represent respectively the values of hue, saturation and intensity of the pixels of coordinates *i,j*.

K_h, K_s, K_i -- constants and calibrated by test.

$\mu_v(H)$, $\mu_v(S)$ and $\mu_v(L)$ --averages of living plants in above-mentioned layers.

$\sigma_v(H)$, $\sigma_v(s)$ and $\sigma_v(L)$--standard deviations of living plants in also above-mentioned layers.

5.Three stages of segmentation of the images

5.1 First stage

After levels of hue, saturation and luminosity are divided in 10 classes, the image is scanned by a mask[4] of L inverted , from left to right and from top to bottom(figure 2).

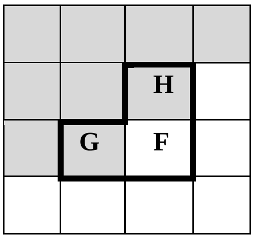


FIG. 2: Mask of inverted L

If mod(x) represents classes (hue, saturation and luminosity) of a pixel x, and Reg(x) the region to which this pixel x belongs, four cases occur :

1.If {mod(F)=mod(G) and mod(F)≠mod(H) }
then Reg(F)=Reg(G)

2. If {mod(F) ≠mod(G) andt mod(F)=mod(H) }
then Reg(F)=Reg(H)

3.If {mod(F) ≠mod(G) and mod(F) ≠mod(H) }
then Reg(F)=K+1, where K+1 is the label of the new region

4.if {mod(F)=mod(G) and mod(F) =mod(H) }
then Reg(F)=Reg(H)=Reg(G) and K=K-1
two regions are merged: Reg(G) and Reg(H)

5.2 Second stage :

When the image has been segmented by first stage, all lengths of frontier of each small region with its neighbors are memorized in a group.

If the region b has N neighbor regions, $l(b,x)$ is the length of frontier between the region b and the region x, and $v(b,x)$, the rate of frontier of the region x such that :

$$v(b,x)=l(b,x)/N$$

The following formula are used to calculate the differences between the region b with its neighbor region:

$$d(b,x) = \sqrt{(H_b - H_x)^2 + (I_b - I_x)^2}$$

H_b, H_x are the averages of the hue of all pixels belonging to region b and region x respectively.

I_b, I_x are the average of the intensity of all pixels belonging to region b and region x respectively .

We give the algorithm of second stage as following:

The region b is merged with its neighbor region x and the label k=k-1 only if :

$v(b,x)=1$

or

$0.5<v(b,x)<1$ *and*
$d(b,x)=min(d(b,x_1),d(b,x_2),d(b,x_3),d(b,x_4),\ldots)$
and $d(b,x)>K_d$

where $x_1, x_2, x_3, x_4,\ldots,$ *are all neighbor regions of the region b*

K_d is constant and calibrated by test.

5.3 The third stage

We give the algorithm of third stage as following:
According to the criterion of recognition ***C.1*** *(mentioned in §3.3.)*
If those small regions can reconstruct a shape of onion or weed leaf and each of those regions has at least a common frontier with another region , they are merged, Otherwise they are not .

6. Results and discussion

Figure 3 and figure 4 show two examples of the results . In these two figures:
(a) -- Original image.
(b) -- Image after discarding non-plant
(c) -- Image after the first stage.
(d) -- Image after the second stage.
(e) -- Image of onion.
(f). -- Image of weed.

All parameters of results (e , f in figure 3 and figure 4) are listed in Table.3 and Table.4.

(a) (b) (c)

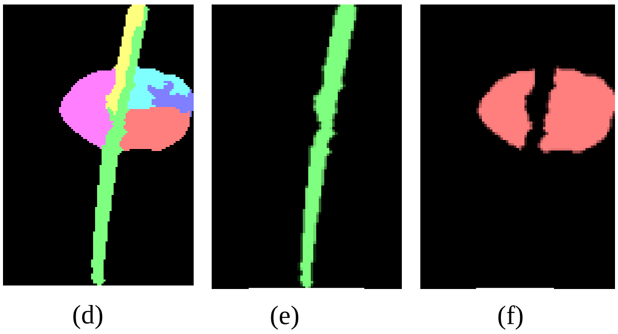


FIG. 3: Images of onion with Sharp-leaved fluellen

TAB. 3: Analysis of the result of figure 3

	Onion	Weed part1	Weed part2
Length	140	39	42
Circularity	81	16	15
Red:	124	153	117
Green	191	219	189
Blue	165	185	127
Hue	0.4159	0.4088	0.3512
Saturation	0.3584	0.3151	0.4193
Intensity	0.7485	0.8603	0.7372
Percentages of sunlit areas	9%	14%	

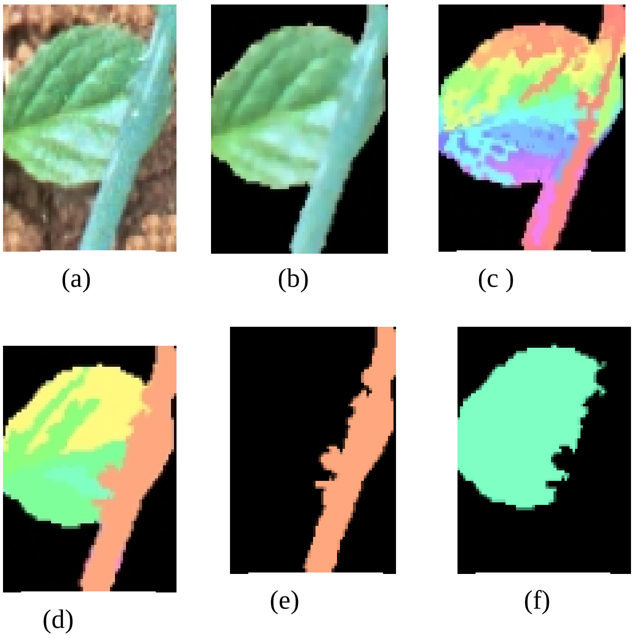


FIG. 4: Images of onion with Sharp- leaved fluellen

TAB. 4: Analysis of the result of figure 4

	Onion	Weed
Length	73	49
Circularity	42	21
Red:	127	129
Green	198	198.
Blue	184	144
Hue	0.4598	0.3657
Saturation	0.3621	0.3719
Intensity	0.7827	0.7775
Percentages of sunlit areas	20%	38%

Everyone can find that their real color components (red, green , blue, hue, saturation and intensity) in the table 3 and table 4 are so approached that confusion is possible between onions and weed leaves if we analyze them only by their color . It is essential to add other attributes like the rates of frontiers and the shape factor of regions. One can also see the criterion of recognition ***C.1*** (mentioned in §3.3.) is able to identify correctly the onion and weeds leaves.

7.Conclusion

The method that we propose here has been verified on several images, we will apply it to the more complex images. In our analysis, the information on the shape has allowed to increase the quality of the identification. Another solution using multi-spectral information will make our method robust.

The weakness of our method is costing a great deal of time for calculating, this will be solved in future by the development of computing technique or modifying our method according to the real conditions.

8. Reference

[1]. J.W.Lu, J. P. Guillemin and X.H.Zhang , " a study for measuring the sunlit areas of onions and weeds In the field by image processing", Proceedings of international conference on engineering and technological sciences 2000(6):230-235.

[2]. J.W.Lu, P.Gouton and J. P. Guillemin, "Utilization of segmentation of color images to distinguish onions and weeds in field". Proceedings of international conference on quality control by artificial vision 2001(2):557-562.

[3] D.E. Guyer, G.E. Miles, L.D. Gaultney, M.M Schreiber, "Application of Machine to Shape analysis in Leaf and Plant identification", Transaction on the ASAE, vol 36(1), pp 163-171,1993.

[4]. P. Gouton**,** "Méthodes et Analyses Des Images Multi-composantes: Application aux images couleur de milieux naturels" . Habilitation à Diriger des Recherches, Université de Bourgogne, Dijon. pp.75-79,2000.