International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.2 (2016), pp.355-368
http://dx.doi.org/10.14257/ijsip.2016.9.2.30

Multi-type Feature Fusion Technique for Weed Identification in Cotton Fields

Guan Lin[1], Liu Zhenzhong[2*], Wu Qiufeng[2], Wang Lulu[1]
[1]*College of Engineering, Northeast Agricultural University*
Harbin 150030, China
[2]*College of Science, Northeast Agricultural University*
Harbin 150030, China
[1]*DNguanlin@163.com;*[2*]*Lzz00@126.com(Correspondence author)*

Abstract

Weed identification is core of precision variable spray technology and weed information management system. Single type features are difficult to identify multi-class weeds in cotton fields. In this paper, multi-type feature fusion technique for weed identification is proposed. Firstly, multi-type features are extracted. In color feature extraction, FMS, SMS and TMS in HSI are extracted by color moment. In shape feature extraction, REC, RWL, CIR and SPH are extracted by geometric parameter method. In texture feature extraction, ASM, CON and COR are extracted by GLCM. Secondly, because feature dimension is too large, principle component analysis is used to reduce dimension to extract new features including COR, ASM, REC and two components. Finally, three comparative experiments including identification of five kinds of weeds, three kinds of weeds and two kinds of weeds are carried out. Experimental results show that method proposed in this paper is superior to state of the art and is suitable for identification of multi-class weeds. This method can also be applied in identifying weeds in other fields.

Keywords: *Weed identification, Multi-type features, Principle component analysis*

1. Introduction

Cotton is one of important economic crops and has important status in development of national economy. Cotton is easily affected by weeds including endives, eclipta prostrate, calystegia hederacea wall, Amaranthus retroflexus L and Amaranthus lividus L *etc*. (See Figure 1). These weeds seriously affect growth and yield of cotton. On the other hand, these weeds rapidly spread and have strong survival ability. Therefore, this paper focuses on identification of endives, eclipta prostrate, calystegia hederacea wall, Amaranthus retroflexus L and Amaranthus lividus L. Weed identification is core of precision variable spray technology and weed information management system.

Weed identification based on computer vision includes image segmentation, feature extraction, dimension reduction and discriminative learning. Feature extraction is the key link to weed identification, which affects the following link and final result of weed identification. Many scholars devoted to feature extraction. Features extracted in weed identification mainly include color features, shape features and texture features. In color features, Patil and Kumar extracted color features of tomato leaves with different diseases in RGB by color moment, and then recognized the diseases of tomato leaves according to color features [1]. Xu extracted color features by percent intensity histogram, percent differential histogram, fourier transform and wavelet packet, and then selected the best features using genetic algorithm to identify nutrient and diseases of tomato [2]. Alamdar

[*] Corresponding Author

ISSN: 2005-4254 IJSIP

and Keyvanpour extracted color features of weeds by quad histogram [3]. In shape features, Swain extracted shape features, and then proposed an automated active shape matching technique based on shape features to identify weeds and crops [4]. Xia extracted shape features of pepper leaves, and then classified pepper leaves by situ detection method based on shape features [5]. Wu extracted shape features of weed edge, which was used to detect weeds in wheat fields [6]. In texture features, Guru extracted texture features by color texture moments, Gray-Level Co-occurrence Matrix (GLCM) and Gabor response, and then used probabilistic neural network to classify flowers [7]. Pydipati extracted texture features in HSI using color co-occurrence matrix, which were used to identify diseased citrus leaves [8].

Figure 1. Cotton fields with weeds. (a) Cotton field with Amaranthus retroflexus L. (b) Cotton field with calystegia hederacea wall. (c) Cotton field with endives. (d) Cotton field with eclipta prostrate. (e) Cotton field with Amaranthus lividus L.

Most of references above focus on using single type features to identify weeds. However, multi-class weeds in cotton fields are difficultly identified by single type features. For example, Amaranthus retroflexus L and Amaranthus lividus L can not be identified by shape features or texture features, but they can be identified by color features; Endives, calystegia hederacea wall and Amaranthus lividus L can not be identified by color features, but they can be identified by shape features or texture features.

In order to effectively identify weeds in cotton fields, this paper proposes multi-type feature fusion technique for weed identification in cotton fields. Effective multi-type features, such as color, shape and texture, are analyzed and extracted in Section 2. In Section 3, dimension of feature parameters is reduced by Principle Component Analysis (PCA). The process of weed identification based on multi-type feature fusion algorithm is given in Section 4. Comparative experiments on weed identification are carried out in Section 5. Conclusion is shown in Section 6.

2. Multi-Type Features Extraction

Since many weeds in cotton fields have similar features in color, shape or texture, it is difficult to identify weeds by single type features. So, this paper extracts multi-type features, such as color features, shape features and texture features, to identify multi-class weeds in cotton fields.

2.1. Comparison and Extraction of Color Features

Because color features are the most intuitive and the most obvious image features, it can effectively identify weeds whose color is obviously different [9]. For example, Amaranthus retroflexus L and Amaranthus lividus L are similar in shape features and texture features (See Figure 2 (a) and (a1)), but can be identified by color features. Not all color features in different color spaces can identify these weeds. For example, R component, G component and B component in RGB can not identify these weeds (See Figure 2 (b)-(d) and (b1)-(d1)), but these weeds can be identified by H component and S component in HSI (See Figure 2 (e)-(f) and (e1)-(f1)). Because these weeds are not accurately identified by H component, this paper only extracts S component.

S component of all pixels in image creates S component feature, which easily causes large dimension of feature. In order to reflect S component feature of image, moments of S component are used. The First Moment of S (FMS), the Second Moment of S (SMS) and the Third Moment of S (TMS) are denoted by

$$FMS = \frac{1}{N}\sum_{j=1}^{N} p_j \tag{1}$$

$$M_k = \left(\frac{1}{N}\sum_{j=1}^{N}\left(p_j - FMS\right)^k\right)^{1/k} \quad \left(k = 2,3\right) \tag{2}$$

where p_j represents S value of j_{th} pixel; N represents the number of pixels in image; M_2 and M_3 represent SMS and TMS respectively.

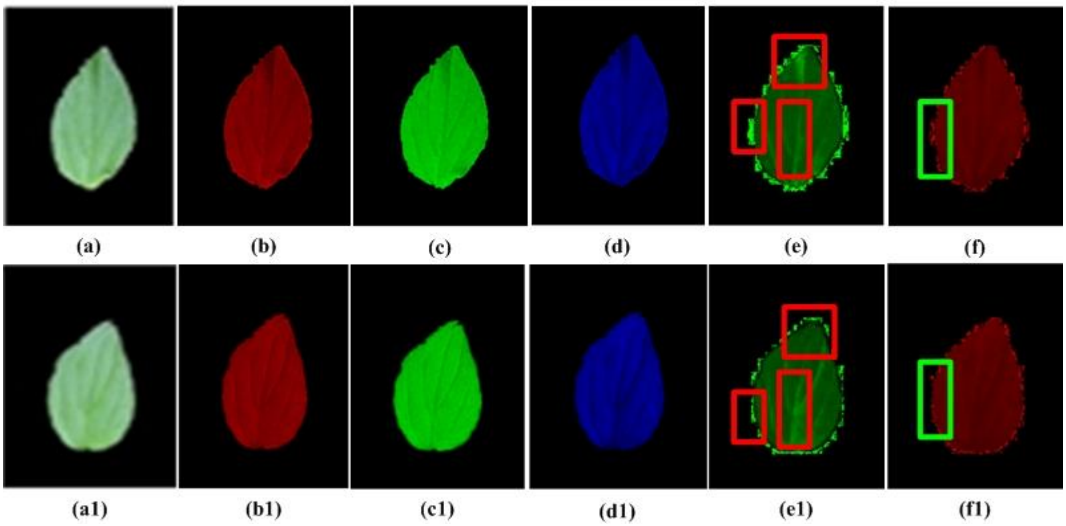


Figure 2. Color features of Amaranthus retroflexus L and Amaranthus lividus L in different color spaces. (a) Original image of Amaranthus retroflexus L. (a1) Original image of Amaranthus lividus L. (b) R component of Amaranthus retroflexus L image in RGB. (b1) R component of Amaranthus lividus L image in RGB. (c) G component of Amaranthus retroflexus L image in RGB. (c1) G component of Amaranthus lividus L image in RGB. (d) B component of Amaranthus retroflexus L image in RGB. (d1) B component of Amaranthus lividus L image in RGB. (e) S component of Amaranthus retroflexus L image in HSI. (e1) S component of Amaranthus lividus L image in HSI. (f) H component of Amaranthus retroflexus L image in HSI. (f1) H component of Amaranthus lividus L image in HSI

In order to verify that moments of S component can be easily used to identify these weeds, an example that identification of Amaranthus retroflexus L and Amaranthus lividus L by moments of S component is given. Moments of S component of Amaranthus retroflexus L and Amaranthus lividus L are obtained by formula (1) and (2) (See Table 1). Table 1 shows that moments of S component, including FMS, SMS and TMS, can easily identify Amaranthus retroflexus L and Amaranthus lividus L.

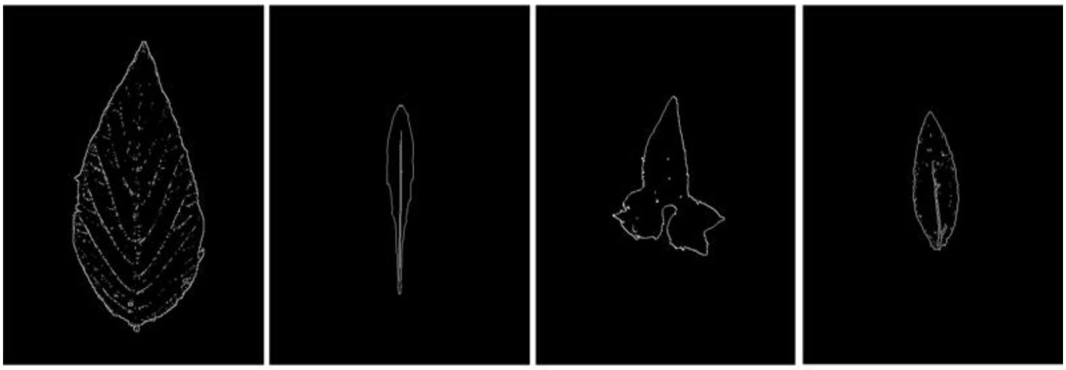
Therefore, this paper extracts color features by color moment and takes FMS, SMS and TMS as color type features.

Table 1.Comparison of FMS, SMS and TMS

Moments	*Weed names*	*Values*
FMS	Amaranthus retroflexus L	0.3835
	Amaranthus lividus L	0.2732
SMS	Amaranthus retroflexus L	0.2374
	Amaranthus lividus L	0.1427
TMS	Amaranthus retroflexus L	0.1393
	Amaranthus lividus L	0.0886

2.2. Comparison and Extraction of Shape Features

Besides color features, shape features are typical features of weed leaves and shape among most weed leaves is different, so it can effectively identify the weeds whose shape is obviously different. For example, shape of Amaranthus retroflexus L leaves, endives leaves, calystegia hederacea wall leaves and eclipta prostrate leaves are elliptic ovate, lanceolate, halberd and ovate respectively (See Figure 3). These differences in shape can be represented by geometric parameters, such as Rectangularity (REC), the Ratio of Width to Length (RWL), Circularity (CIR) and Sphericity (SPH). Meaning and formulas of these geometric parameters are given in Table 2.



(a) (b) (c) (d)

Figure 3. Shape of Different Weed Leaves. (a) Shape of Amaranthus Lividus L leaf. (b) Shape of Endives Leaf. (c) Shape of Calystegia Hederacea Wall Leaf. (d) Shape of Eclipta Prostrate Leaf

Table 2. Geometric Parameters

Feature parameters	*Meaning of parameters*	*Calculation formulas*
REC	Degree of minimum circumscribed rectangle filled by leaf area	R=A/(L*W)
RWL	Ratio of width to length of minimum circumscribed rectangle	K=W/L
CIR	Correlation degree of leaf area and minimum circumscribed circle	D=4πA/L2
SPH	Ratio of leaf area to perimeter of minimum circumscribed rectangle	F=A/(2L+2W)

In Table 2, A, W and L represent the number of pixels in the leaf area, width of minimum circumscribed rectangle and length of minimum circumscribed rectangle respectively.

In order to verify that weeds can be identified by geometric parameters, identification of weeds above is taken as an example. Geometric parameters, including REC, RWL, CIR and SPH, are obtained by formulas in Table 2 (See Table 3). From Table 3, geometric parameters of different weeds are obviously different.

Table 3. Geometric Parameters of Weeds

Shape features / *Weed names*	*CIR*	*SPH*	*REC*	*RWL*
Amaranthus lividus L	3.8419	45.6020	0.6111	1.9955
Endives	2.1900	13.8296	0.5726	3.2857
Calystegia hederacea wall	3.9740	23.6633	0.4294	1.3316
Eclipta prostrate	3.3831	33.5634	0.7179	2.6667

In Table 3, the values of geometric parameters are obtained by formulas in Table 2 with invariance of image rotation, that is, rotation can not influence shape feature extraction.

Therefore, this paper extracts geometric parameters, including REC, RWL, CIR and SPH, as shape type features.

2.3. Comparison and Extraction of Texture Features

Besides intuition of color features and typicality of shape features, texture features can reflect microscopic information of images and centralize choroid of weed leaves [10]. Since choroid of most weed leaves is different, texture features can effectively identify weeds. For example, choroid of Amaranthus lividus L leaf, endives leaf, calystegia hederacea wall leaf and eclipta prostrate leaf are relatively shallow, sparse, rough and complex respectively (See Figure 4). It is difficult to describe differences of texture features of images by direct comparison of choroid, but GLCM can describe these differences.

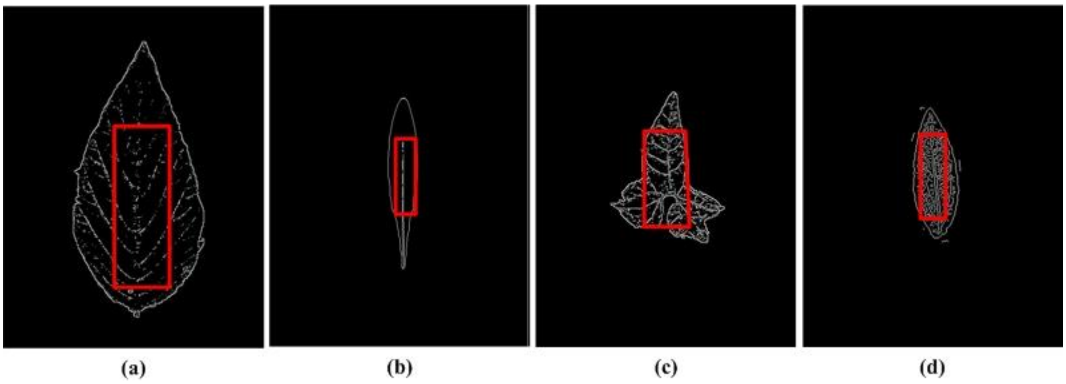


Figure 4. Choroid of Different Weed Leaves. (a) Choroid of Amaranthus lividus L leaf (b) Choroid of Endives Leaf. (c) Choroid of Calystegia Hederacea Wall Leaf. (d) Choroid of Eclipta Prostrate Leaf

GLCM is proposed by Haralick. GLCM is denoted by

$$p_{d,\theta}(i,j) = counts\left\{\left\{(x_1,y_1),(x_2,y_2)\right\} \in M \times N \mid f(x_1,y_1)=i, f(x_2,y_2)=j\right\} \quad (3)$$

where $(x1, y1)$ and $(x2, y2)$ are pixels in a $M \times N$ gray image respectively; i and j represent gray level of $(x1, y1)$ and $(x2, y2)$ respectively; d is the distance between $(x1, y1)$ and $(x2, y2)$; θ is angle between the connection of two pixels and the horizontal axis. Usually, θ takes 0°, 45°, 90° and 135°.

In GLCM, 14 parameters were defined, but only four parameters of GLCM including Angular Second Moment (ASM), Contrast (CON), Correlation (COR) and Entropy (ENT) are verified that their correlation each other is small by experiments [11]. The formulas of ASM, CON, COR and ENT are given in Table 4.

Table 4. The formulas of ASM, CON, COR and ENT

Parameters	*Calculation formulas*	
ASM	$f_1 = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} p_d^2(i,j)$	$\mu_1 = \sum_{i=0}^{L-1} i \sum_{j=0}^{L-1} p_d(i,j)$ $\mu_2 = \sum_{i=0}^{L-1} j \sum_{j=0}^{L-1} p_d(i,j)$ $\sigma_1^2 = \sum_{i=0}^{L-1}(i-\mu_1)\sum_{j=0}^{L-1} p_d(i,j)$ $\sigma_2^2 = \sum_{i=0}^{L-1}(j-\mu_2)\sum_{j=0}^{L-1} p_d(i,j)$
CON	$f_2 = \sum_{n=0}^{L-1} n^2 \left\{\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} p_d(i,j)\right\}$	
COR	$f_3 = \frac{\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} ijp_d(i,j) - \mu_1\mu_1}{\sigma_1^2\sigma_2^2}$	
ENT	$f_4 = -\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} p_d(i,j)\log p_d(i,j)$	

Texture features of weeds, including Amaranthus lividus L, endives, calystegia hederacea wall and eclipta prostrate, are taken as an example. Texture features of these weeds are extracted by GLCM (See Table 5). Table 5 shows that texture features of these weeds are obviously different except ENT.

Therefore, this paper extracts texture features by GLCM and takes ASM, CON and COR as texture type features.

Table 5. Texture Features Eextracted by GLCM

Texture features / *Weed name*	*ASM*	*CON*	*COR*	*ENT*
Amaranthus lividus L	0.5613	1.1077	0.0042	0.0046
Endives	0.8122	0.6034	0.0095	0.0041
Calystegia hederacea wall	0.8479	0.5042	0.0121	0.0041
Eclipta prostrate	0.9077	0.0204	0.0086	0.0040

In summary, methods, color spaces and feature parameters in single type feature extraction are shown in Table 6. In color feature extraction, FMS, SMS and TMS in HSI are extracted by color moment. In shape feature extraction, REC, RWL, CIR and SPH are extracted by geometric parameter method. In texture feature extraction, ASM, CON and COR are extracted by GLCM. All features above constitute features space of weed identification.

Table 6. Methods, Color Spaces and Feature Parameters of Single Type Features

Methods	*Color spaces*	*Feature parameters*
Color moment	HSI	FMS, SMS, TMS
Geometric parameter method	Gray	REC, RWL, CIR, SPH
GLCM	Gray	ASM, CON, COR

3. Dimension Reduction of Feature Parameters

These feature parameters, including FMS, SMS, TMS, REC, RWL, CIR, SPH, ASM, CON and COR, have a certain correlation, which not only increases complexity and computation of feature extraction, but also causes information overlap. So, this paper reduces dimension of features by PCA.

Correlation coefficient matrix of these features is obtained according to dataset observed. In Table 7, correlation coefficients between COR, ASM and REC and others are less than 0.7, which are considered that correlation is smaller, so COR, ASM and REC are directly kept. Other feature parameters whose correlation coefficients are more than 0.7 are indirectly represented by principle components [12].

Table 7. The Correlation Coefficient Matrix

	FMS	*SMS*	*TMS*	*REC*	*RWL*	*CIR*	*SPH*	*ASM*	*CON*	*COR*
FMS	1									
SMS	0.854	1								
TMS	0.832	0.873	1							
REC	0.476	0.513	0.437	1						
RWL	0.859	0.909	0.854	0.456	1					
CIR	-0.704	-0.767	-0.665	-0.68	-0.81	1				
SPH	-0.759	-0.804	-0.726	-0.66	-0.841	0.95	1			
ASM	0.102	0.135	0.037	0.559	0.083	-0.5	-0.504	1		
CON	0.668	0.703	0.622	0.69	0.705	-0.8	-0.824	0.506	1	
COR	0.254	0.279	0.222	0.672	0.195	-0.5	-0.5	0.66	0.618	1

Two principle components are extracted. The first principle component accounts for 66.534% of the total standardized variance and the second principle component accounts for 19.108% (See Table 8). Two principal components are created by seven feature parameters (See Table 9) and are taken as features to identify weeds.

Table 8. Total Variance Explained by PCA

Principle components	*Initial eigenvalues*		
	Total	*%of Variance*	*%of Cumulative*
1	6.653	66.534	66.534
2	1.911	19.108	85.642

Table 9. Loading Matrix

	Principle components	
	1	2
FMS	0.844	-0.371
SMS	0.879	-0.363
TMS	0.813	-0.438
RWL	0.876	-0.419
CIR	-0.928	-0.078
SPH	-0.948	-0.006
CON	0.896	0.164

Based on these, feature dimension is reduced from 10 to 5 and 5 features, including COR, ASM, REC and two principle components, are obtained to identify these weeds.

4. Weed Identification Based on Multi-Type Feature Fusion

In summary, steps of weed identification based on multi-type feature fusion algorithm are elaborated by:

Step 1: Create training dataset and testing dataset. 40 images of each class in these weeds, such as endives, eclipta prostrate, calystegia hederacea wall, Amaranthus retroflexus L and Amaranthus lividus L, are selected as training dataset and 20 images of each class in these weeds are selected as test dataset.

Step 2: Extraction of multi-type features. In color feature extraction, FMS, SMS and TMS in HSI are extracted by color moment. In shape feature extraction, REC, RWL, CIR and SPH are extracted by geometric parameter method. In texture feature extraction, ASM, CON and COR are extracted by GLCM.

Step 3: Dimension reduction of feature parameters. Feature parameters whose correlation coefficients with others are less than 0.7 are directly kept, including COR, ASM and REC. Other feature parameters are indirectly represented by two principle components, which are as new features with COR, ASM and REC.

Step 4: Weeds in testing dataset are identified by k-nearest neighbor.

5. Comparative Experiment

In order to verify accuracy of weed identification in cotton fields, three comparative experiments are carried out. Three experiments include identification of five kinds of weeds, three kinds of weeds and two kinds of weeds.

In order to express clearly, names of weeds and accuracy of identification by different methods are represented by abbreviations (See Table 10).

Table 10. Full Names and Abbreviations

Full names	*Abbreviations*
Accuracy of weed identification by color features	CF
Accuracy of weed identification by shape features	SF
Accuracy of weed identification by texture features	TF
Accuracy of weed identification by multi-type features	AF
Accuracy of weed identification by principle components that directly extracted by PCA	DPCA
Accuracy of weed identification by features that extracted by method proposed in this paper.	IPCA
Amaranthus retroflexus L	I
Calystegia hederacea wall	II
Endives	III
Eclipta prostrate	IV
Amaranthus lividus L	V

5.1. Comparative Experiment on Five kinds of Weeds

This experiment on five kinds of weeds, including endives, eclipta prostrate, calystegia hederacea wall, Amaranthus retroflexus L and Amaranthus lividus L, is carried out. Accuracy of weed identification in each class and total accuracy of identification by different methods are shown in Table 11.

Table 11. Accuracy of Identification by Different Methods for Five Kinds of Weeds

	I	*II*	*III*	*IV*	*V*	*Total*
CF	44%	51.5%	40%	54%	55.5%	45%
SF	77.5%	77.5%	72.5%	65%	62.5%	71%
TF	70%	75%	72.5%	70%	67.5%	69%
AF	87.5%	85%	85%	82.5%	80%	84%

Table 11. Accuracy of Identification by Different Methods for Five Kinds of Weeds (Continued)

	I	*II*	*III*	*IV*	*V*	*Total*
DPCA	88%	85.5%	86%	84%	81%	84.9%
IPCA	89%	87%	87.5%	87%	86.5%	88%

From Table 11, accuracy of identification by single type features is significantly lower than accuracy of identification by multi-type features. The experimental results show that accuracy of identification by IPCA is 88% which is higher than others.

5.2. Comparative Experiment on Two Kinds of Weeds

This experiment on any two kinds of weeds among endives, eclipta prostrate, calystegia hederacea wall, Amaranthus retroflexus L and Amaranthus lividus L, is carried out. Accuracy of weed identification in each class and total accuracy of identification by different methods are shown in Table 12.

Table 12. Accuracy of Identification by Different Methods for Two Kinds of Weeds

		CF	*SF*	*TF*	*AF*	*DPCA*	*IPCA*
Two kinds of weeds	I	55%	65%	68.75%	80%	82.5%	85%
	II	57.5%	75%	75%	85%	82.5%	83%
	Total	56.25%	70%	71.25%	82.5%	82.5%	**84%**
	I	52.5%	72.5%	67.5%	77.5%	80%	82%
	III	45%	75%	72.5%	82.5%	85%	86%
	Total	48.75%	73.75%	70%	80%	82.5%	**84%**
	I	55%	70%	67.5%	77.5%	82.5%	81.25%
	IV	60%	75%	77.5%	87.5%	87.5%	90%
	Total	57.5%	72.5%	72.5%	82.5%	85%	**86.25%**
	I	52%	70%	70%	77.5%	80%	82.5%
	V	48%	77.5%	75%	82.50%	85%	87.5%
	Total	49.25%	73.75%	72.5%	80%	82.5%	**85%**
	II	55%	72.5%	72.5%	75%	77.5%	85%
	III	50%	67.5%	70%	80%	82.5%	82.5%
	Total	52.5%	70%	71.25%	77.5%	80%	**83.75%**
	II	55.5%	75%	67.5%	75%	77.5%	80%
	IV	57.5%	67.5%	72.5%	85%	87.5%	90%
	Total	56.5%	71.25%	70%	80%	82.50%	**85%**

	Ⅱ	57.5%	75%	75%	77.50%	82.50%	82.5%
	Ⅴ	55%	70%	67.5%	80%	85%	87.5%
	Total	56.25%	72.5%	71.25%	78.75%	83.75%	**85%**
	Ⅲ	57.5%	72.5%	72.5%	77.5%	82.5%	85%
	Ⅳ	52.5%	70%	65%	75%	80%	82.5%
	Total	55%	71.25%	68.75%	76.5%	81.25%	**83.75%**

Table 12. Accuracy of Identification by Different Methods for Two Kinds of Weeds (Continued)

Two kinds of weeds	Ⅲ	50%	72.5%	70%	82.5%	85%	87.5%
	Ⅴ	48%	67.5%	67.5%	75%	77.5%	85%
	Total	49%	70%	68.75%	78.75%	81.25%	**86.25%**
	Ⅳ	52.5%	72.5%	70%	82.5%	85%	90%
	Ⅴ	48%	65%	65%	77.5%	80%	82.5%
	Total	50.25%	68.75%	67.5%	80%	82.50%	**86.25%**

From Table 12, the highest accuracy of identification by CF, SF, TF, AF, DPCA and IPCA are 56.25%, 73.75%, 72.5%, 82.5%, 85% and 86.25% respectively and accuracy of identification by IPCA is higher than that by others for any two kinds of weeds.

5.3. Comparative experiment on three kinds of weeds

This experiment on any three kinds of weeds among endives, eclipta prostrate, calystegia hederacea wall, Amaranthus retroflexus L and Amaranthus lividus L, is carried out. Accuracy of weed identification in each class and total accuracy of identification by different methods are shown in Table 13.

From Table 13, the highest accuracy of identification by CF, SF, TF, AF, DPCA and IPCA are 46.7%, 78.3%, 72.5%, 82.5%, 85% and 87.5% respectively and accuracy of identification by IPCA is higher than that by other methods for any three kinds of weeds.

Table 13. Accuracy of Identification by Different Methods for Three Kinds of Weeds

		CF	*SF*	*TF*	*AF*	*DPCA*	*IPCA*
Three kinds of weeds	Ⅰ	40%	77.5%	67.5%	77.5%	82.5%	90%
	Ⅱ	42.5%	75%	70%	85%	82.5%	85%
	Ⅲ	45%	77.5%	75%	82.5%	85%	87.5%
	Total	42.5%	76.7%	70.8%	81.67%	83.33%	**87.5%**
	Ⅰ	40%	77.5%	72.5%	75%	80%	82.5%
	Ⅱ	42.5%	77.5%	67.5%	82.5%	82.5%	87.5%
	Ⅳ	47.5%	80%	75%	82.5%	87.5%	90%
	Total	45%	78.3%	71.7%	80%	84.2%	**86.7%**
	Ⅰ	40%	75%	67.5%	85%	87.5%	84.2%
	Ⅱ	45%	77.5%	70%	75%	85%	88.3%
	Ⅴ	42.5%	72.5%	75%	82.5%	80%	85%
	Total	43.3%	76.7%	72.5%	80%	82.5%	**86.7%**
	Ⅰ	35%	77.5%	65%	77.5%	82.5%	85%

	Ⅲ	47.5%	72.5%	70%	85%	80%	88.3%
	Ⅳ	45%	70%	72.5%	82.5%	87.5%	84.2%
	Total	42.5%	75%	70%	81.7%	83.33%	**86.7%**
	I	35%	70%	65%	87.5%	85%	85%
	Ⅲ	47.5%	82.5%	70%	75%	77.5%	88.3%
	V	37.5%	80%	77.5%	85%	87.5%	92.5%
	Total	40.8%	77.5%	70.8%	82.5%	84.2%	**86.7%**

Table 13. Accuracy of Identification by Different Methods for Three Kinds of Weeds (Continued)

		CF	*SF*	*TF*	*AF*	*DPCA*	*IPCA*
Three kinds of weeds	I	42.5%	82.5%	65%	82.5%	82.5%	87.5%
	Ⅳ	37.5%	70%	67.5%	75%	80%	85%
	V	52.5%	72.5%	75%	82.5%	87.5%	82.5%
	Total	44.1%	75.8%	70%	80%	83.3%	**85%**
	Ⅱ	52.5%	70%	67.5%	75%	82.5%	95%
	Ⅲ	45%	82.5%	70%	85%	87.5%	90%
	Ⅳ	40%	80%	75%	82.5%	85%	87.5%
	Total	45.8%	77.5%	72.5%	82.5%	85%	**87.5%**
	Ⅱ	50%	75%	65%	77.5%	77.5%	82.5%
	Ⅲ	40%	77.5%	72.5%	80%	82.5%	85%
	V	37.5%	82.5%	70%	85%	85%	92.5%
	Total	42.5%	78.3%	71.7%	80.8%	81.7%	**87.5%**
	Ⅱ	40%	70%	62.5%	75%	75%	82.5%
	Ⅳ	55%	75%	70%	80%	82.5%	85%
	V	45%	82.5%	75%	82.5%	85%	87.5%
	Total	46.7%	75.8%	70%	79.2%	80%	**85%**
	Ⅲ	35%	75%	62.5%	85%	87.5%	88.3%
	Ⅳ	47.5%	80%	67.5%	72.5%	80%	87.5%
	V	37.5%	70%	77.5%	82.5%	85%	85%
	Total	40.8%	75%	69.2%	80%	82.5%	**86.7%**

By integrating experiments above, accuracy of identification by IPCA is higher than other methods, such as CF, SF, TF, AF and DPCA, for any kinds of weeds. According to the highest accuracy of identification and total accuracy of identification in different comparative experiments, IPCA is most suitable for the identification of five kinds of weeds, and then is more suitable for three kinds of weeds, which shows that IPCA is suitable for multi-class weed identification.

6. Conclusion

This paper proposes multi-type feature fusion technique, which can solve the problem of identification of multi-class weeds in cotton fields. Firstly, Multi-type features are extracted. In color feature extraction, FMS, SMS and TMS in HSI are extracted by color moment. In shape feature extraction, REC, RWL, CIR and SPH are extracted by geometric parameter method. In texture feature extraction, ASM, CON and COR are extracted by GLCM. Secondly, dimension of feature parameters is reduced by PCA. Feature dimension is reduced from 10 to 5 and 5 features, including COR, ASM, REC and two principle components, are regarded as new features to identify weeds. Finally, three comparative experiments are carried out. Accuracy of identification by IPCA for five kinds of weeds, three kinds of weeds and two kinds of weeds are 88%, 87.5% and 86.25% respectively, which is higher than state of the art, so IPCA is more suitable for identification of multi-class weeds. This method can be also used for weed identification in other fields.

Acknowledgments

This work was supported in part by the Nature Science Foundation of Northeast Agricultural University under contract no.2011RCA01 and National Key Technology R&D Program of the Ministry of Science and Technology under contract 2014BAD-12B01.

References

[1] [1] Jayamala K.Patil and Raj Kumar, “Color Feature Extraction of Tomato Leaf Diseases”, International Journal of Engineering Trends and Technology, vol. 2, no.2, (2011), pp.72-74.
[2] [2] Guili Xu, Fengling Zhang and Syed Ghafoor Shah, “Use of Leaf Color Images to Identify Nitrogen and Potassium Deficient Tomatoes”, Pattern Recognition Letters, vol. 32, no.11, (2011), pp. 1584-1590.
[3] [3] Fatemeh Alamdar and MohammadReza Keyvanpour, “A New Color Feature Extraction Method Based on Quad Histogram”, Procedia Environmental Sciences, Vol. 10, (2011), pp. 777-783.
[4] [4] Kishore C. Swain, Michael Nørremark, Rasmus N. Jørgensen, “Weed Identification Using an Automated Active Shape Matching Technique”, Biosystems Engineering, vol. 110, no.4, (2011), pp. 450-457.
[5] [5] Chunlei Xia, Jang-Myung Lee and Yan Li, “Plant Leaf Detection Using Modified Active Shape Models,” Biosystems Engineering, vol. 116, no. 1, (2013), pp. 23-35.
[6] [6] Xuewen Wu, Wen qiang Xu, Yunyun Song and Mingxing Cai, “A Detection Method of Weed in Wheat Field on Machine Vision”, Procedia Engineering, vol. 15, (2011), pp. 1998-2003.
[7] [7] TD.S.Guru, Y.H.SharathKumar and S.Manjunath, “Textural Features in Flower Classification”, Mathematical and Computer Modeling, vol. 54, no. 3–4, (2011), pp. 1030-1036.
[8] [8] R. Pydipati, T.F. Burks and W.S. Lee, “Identification of Citrus Disease Using Color Texture Features and Discriminant Analysis”, Computers and Electronics in Agriculture, Vol. 52, No. 1–2, (2011), pp. 49-59.
[9] [9] Huihui Wang, Yang Lv, Yonghai Sun, Xueheng Tao and Jixin Yang, “Image Color and Texture-based Modeling and Simulation of Ripeness Identification for Fresh Corn Ears”, International Journal of Processing, Image Processing and Pattern Recognition, vol.6, no.6, (2013), pp. 341-350.
[10] [10] Jun Yue, Zhenbo Li and Lu Liu, Zetian Fu, “Content-based Image Retrieval Using Color and Texture Fused Features”, Mathematical and Computer Modelling, vol.54, (2011), pp. 1121–1127.
[11] [11] Jagadeesh D.Pujari, Rajesh Yakkundimath and Abdulmunaf S.Byadgi, “Classification of Fungal Disease Symptoms Affected on Cereals Using Color Texture Features”, International Journal of Processing, Image Processing and Pattern Recognition, vol.6, no.6, (2013), pp. 321-330.
[12] [12] Mahmood R Golzarian and Ross A Frick, “Classification of Images of Wheat, Ryegrass and Brome Grass Species at Early Growth Stages Using Principal Component Analysis”, Golzarian and Frick Plant Methods, vol.7, (2011), pp.7-28.

Authors

Guan Lin, he received the B.S. degree in Management Science and Engineering Institute, Shandong Institute of Business and Technology. Now he is pursing his M.S. degree in the College of Engineering, Northeast Agriculture University. His research work focuses on pattern recognition and optimization algorithm.

Liu Zhenzhong, he is working as Professor in the College of Science, Northeast Agriculture University. He has more than 28 years of experience in teaching. His research direction is Multivariate Statistical and Biomathematics.

Wu Qiufeng, he was born at Heilongjiang Province, China, in 1979. He received Ph. D in computer application technology from Harbin Institute of Technology, Harbin, China, in January 2014. From July 2002 to the present, he is working in College of Science in Northeast Agricultural University. He is CCF member. His current research interests include image restoration, discriminative learning, and computer vision.

Wang Lulu, he received the B.S. degree in Management Science and Engineering Institute, Shandong Institute of Business and Technology. Now he is pursing his M.S. degree in the College of Engineering, Northeast Agriculture University. His research work focuses on image pattern recognition and optimization algorithm.

