The 9ª International Conference on Advances in Sensors, Actuators, Metering and Sensing
STAAS: Advanced Sensors and Actuators for Agriculture and Knowledge in Engineering

Dimensionality Reduction for CCD Sensor-Based Image to Control Fall Armyworm in Agriculture

Alex B. Bertolla and Paulo E. Cruvinel

Embrapa Instrumentation, São Carlos, SP, Brazil
Post-Graduation Program in Computer Science - Federal University of São Carlos, SP, Brazil

Summary

Abstract

This article presents a study on the **dimensionality reduction** of features from a digital image acquired with a **Charge-Coupled Devices** sensor in an agricultural field, in order to choose the optimal number of **principal components** for reducing feature dimensionality. In this context, it has become very important to define a method for selecting the optimal number of principal components for dimensionality reduction, while retaining only the necessary information associated with the main variables that describe the object of interest, the Fall armyworms (*Spodoptera frugiperda*). The results showed using **Hu invariant moments** for feature extraction, dimensionality reduction was possible for all analyzed cases, **preserving the semantic characteristics** collected by the sensor and **prepare them for classification**.

Introduction

- Charge-Coupled-Devices (CCD)
  - Charge-Coupled-Devices are the most used imaging sensors for digital image acquisition
  - In agriculture, those sensors are usually used to capture images of pests and plant diseases

- The Fall Armyworm (Spodoptera frugiperda)
  - Pests attacks affect the plant production: Economy affected
  - Conventional method to identify pests: Executed by humans and manually
    - Exhaustive activity and passive of errors

The Fall Armyworm

Health cob

Cob under pest attack

Introduction

- Dimensionality Reduction
  - Complex and high dimensions of the data captured
  - Storing and processing the amount of data acquired has become a challenging task
  - Transform the original high-dimensional data into a new reduced dataset
  - Avoid overfitting in classification models

- Proposal
  - Dimensionality reduction optimization when using a CCD sensor-based images to control Fall armyworms in agriculture
  - Hu invariant moments descriptor for geometrical feature extraction
  - The use of Principal Component Analysis(PCA) for feature vector reduction

Material and Methods

- Digital Image Processing and the Dataset
    - A digital image can be defined as a bi-dimensional function *f(x, y)*, where *(x, y)* are the intensity positions, defined as pixel
    - CCD's sensors can capture images in different color spaces, however, the most common color space is the Red, Green, and Blue (RGB), which represents the visible spectrum
    - Regarding the image acquisition, a dataset was generated using a CCD sensor. This dataset is composed of the Fall armyworm images in real maize crops, where the pest was found both in leaves and cobs maize

Image type	JPG / JPEG
Color space	RGB
Width	3072 pixels
Height	2048 pixels
Resolution	72 pixels per inch (ppi)
Pixel size	0,35mm

Image features acquired by CCD sensor

Material and Methods

- Feature Extraction
    - The seven Hu invariant moments descriptor have been considered for the extraction of the geometric features of the pest
    - The first moment is analogous to the moment of inertia of the image and is defined as following

$$\phi_1 = \eta_{20} + \eta_{02}$$

    - The second moment is a measure of the asymmetry of the image

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

Material and Methods

- Feature Extraction
  - The third moment is a measure of the curvature of the image

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

  - The fourth is a measure of the concentration of the image's mass around the centroid

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

Material and Methods

- Feature Extraction
  - The fifth is a measure of the roughness of the image

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \\ [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03}) \\ (\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

  - The sixth is a measure of the asymmetry of the image with respect to the horizontal axis

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

Material and Methods

- Feature Extraction
  - The seventh is a measure of the asymmetry of the image with respect to the vertical axis

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})$$
$$[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]+$$
$$(3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

Material and Methods

- Principal Components Analysis (PCA)
  - PCA considers an array *X* of data with *n* samples representing the number of observations and m independent variables

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix}$$

  - The principal components are obtained for a set of *m* variables *X1, X2, ..., Xm* with means *μ1, μ2, ..., μm* and variance *σ2 1, σ2 2, ..., σ2m*, which are independent and have covariance between the n-th and m-th variable, where *Σ* represents the covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11}^2 & \cdots & \sigma_{1m}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{n1}^2 & \cdots & \sigma_{nm}^2 \end{bmatrix}$$

Material and Methods

- Principal Components Analysis (PCA)
  - The pairs of eigenvalues and eigenvectors are found ($\lambda 1, e1$), ($\lambda 2, e2$), ..., ($\lambda m, em$), where $\lambda 1 \geq \lambda 2 \geq \ldots \geq \lambda m$ and associated with Σ, where the *i-th* principal component is defined by:

$$Z_i = e_{i1}X_1 + e_{i2}X_2 + \ldots + e_{im}X_m$$

  - The objective is to maximize the variance of Zi *where* $i = 1, \ldots, m$, as:

$$Var(Z_i) = Var(e_i'\mathbf{X}) = e_i'Var(\mathbf{X})e_i = e_i'\mathbf{\Sigma}e_i$$

  - The spectral decomposition of the matrix Σ is given by Σ = PΛP′, where P is the composite matrix by the eigenvectors of Σ, and Λ the diagonal matrix of eigenvalues of Σ

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{bmatrix}$$

Material and Methods

- Principal Components Analysis (PCA)
    - The principal component of greatest importance is defined as the one with the greatest variance which explains the maximum variability in the data vector
    - The second most important component is the component that has the second highest variance and so on
    - The normalized eigenvectors represent the main components that constitute the feature vector with reduced dimension
    - The reduced features are used for the recognition of the patterns of Fall armyworm (*Spodoptera frugiperda*)

Results and Discussions

- Image Dataset
    - The image dataset represents the Fall armyworm (*Spodoptera frugperda*) acquired in a real environment of maize crop in its five different stages of growth.
    - Grouped in 456 images for each stage.

Results and Discussions

- Feature Extraction
  - Hu invariant moments and the feature vector normalization.

Hu invariant moments	Images		
	Image 1	**Image 2**	**Image 3**
ϕ_1	6.692	6.6178	6.524
ϕ_2	13.581	13.424	19.102
ϕ_3	24.321	23.944	22.370
ϕ_4	25.919	26.245	23.445
ϕ_5	51.517	-52.023	46.665
ϕ_6	34.307	-33.305	-34.728
ϕ_7	-51.458	-51.556	47.656

The 7 Hu invariant moments

Hu invariant moments	Images		
	Image 1	**Image 2**	**Image 3**
ϕ_1	0.274	0.162	0.021
ϕ_2	-1.048	-1.121	1.496
ϕ_3	1.035	0.817	-0.092
ϕ_4	1.408	1.669	-0.575
ϕ_5	0.719	-1.607	0.610
ϕ_6	0.808	-1.289	-1.333
ϕ_7	-0.863	-0.865	1.199

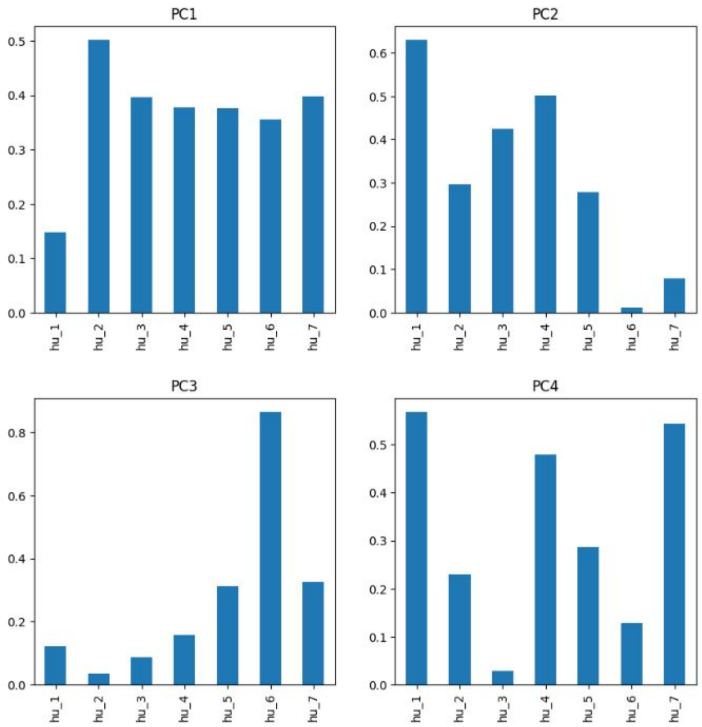
The normalized values

Results and Discussions

- Feature Vector Dimensionality Reduction
    - Principal Component Analysis

Hu invariant moments	Principal components			
	PC 1	PC 2	PC 3	PC 4
ϕ_1	-0.147	-0.630	-0.120	-0.568
ϕ_2	0.501	-0.295	0.036	0.230
ϕ_3	-0.395	-0.424	0.087	0.027
ϕ_4	-0.378	0.501	-0.158	-0.479
ϕ_5	-0.376	-0.277	-0.310	0.286
ϕ_6	-0.355	-0.011	0.865	0.127
ϕ_7	0.398	-0.078	0.325	-0.543

Maximum variation

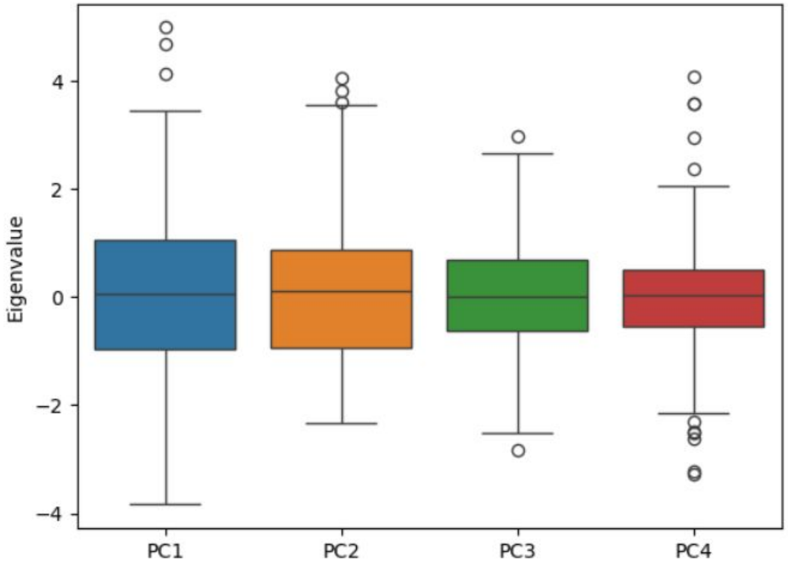


Graphical representation of the maximum variation

Results and Discussions

- Feature Vector Dimensionality Reduction
  - Principal Component Analysis

Principal components	Images		
	Image 1	**Image 2**	**Image 3**
PC1	0.333	2.121	-0.742
PC2	-2.280	-1.156	1.306
PC3	0.551	-1.243	-0.528
PC4	0.181	-1.193	0.012



Conclusion

A study of dimensionality reduction using Principal Components Analysis (PCA) was presented, considering feature vectors composed of Hu invariant moments extracted from digital images acquired with the CCD’s sensors. The measure of the explained variance ratio to the original data was applied to verify the quantity number of principal components necessary to explain the maximum of the original data. The first and fourth invariant moments were used to infer the estimated size of the Fall armyworm (*Spodoptera frugperda*). The measurements showed that computing two to four principal components was sufficient to explain 55% to 80% of the original data. Finally, despite seven invariant moments being used, such analysis led to the conclusion that when using four principal components, one may achieve the explanation of 80% for the original data, with low error, as well as, not a significative variation. For future works, it is suggested to extend this research to an unsupervised method to reach the selection of the number of principal components to remain with the semantic features from a recognized agricultural pest.

Acknowledgment

This research was partially supported by the São Paulo Research Foundation (FAPESP 17/19350-2). We thank the Brazilian Corporation for Agricultural Research (Embrapa) and the Postgraduate Program in Computer Science from the Federal University of São Carlos (UFSCar). The Authors also recognize the helpful discussions with MSc Bruno M. Moreno to finalize the manuscript.

THANK YOU

E-mails: alex.bertolla@embrapa.br; paulo.cruvinel@embrapa.br

Dimensionality Reduction for CCD Sensor-Based Image to Control Fall Armyworm in Agriculture

Alex B. Bertolla and Paulo E. Cruvinel

Embrapa Instrumentation, São Carlos, SP, Brazil
Post-Graduation Program in Computer Science - Federal University of São Carlos, SP, Brazil

FAPESP Embrapa ufscar CAPES