Automatic Fruit and Vegetable Recognition Based on CENTRIST and Color Representation

Jadisha Yarif Ramírez Cornejo and Helio Pedrini[✉]

Institute of Computing, University of Campinas, Campinas, SP 13083-852, Brazil
helio@ic.unicamp.br

Abstract. Automatic fruit and vegetable recognition from images is still a very challenging task. In this work, we describe and analyze an efficient and accurate fruit and vegetable recognition system based on fusing two visual descriptors: Census Transform Histogram (CENTRIST) and Hue-Saturation (HS) Histogram representation. Initially, background subtraction is applied to the fruit and vegetable images. CENTRIST and HS-Histogram are extracted, as well as Color CENTRIST features for comparison purpose. Then, the feature vector is reduced through Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). In the recognition process, K-nearest neighbor (K-NN) and Support Vector Machine (SVM) classifiers are employed. Experiments conducted on a benchmark demonstrate that combining CENTRIST and HS-Histogram representation reached high and competitive recognition accuracy rates compared to other similar works in the literature.

Keywords: Fruit and vegetable recognition · CENTRIST · Color features

1 Introduction

The development of effective recognition systems has many practical applications, however, it is still a challenging task. One of the main problems remains in determining whether or not the selected data contains representative and discriminative information for the classes under consideration. A challenge is to achieve comparable accuracy rates to human recognition.

Fruit and vegetable recognition methods can be applied to crop monitoring, supermarket checkout systems, educational purposes, among others [1,5,9]. Accurate and efficient identification of different species of fruits and vegetables is essential in supermarkets, which must invest time and money in training qualified cashiers. Possible solutions to this problem include the use of price tags or barcodes on each item of the products. When a price change is required, barcodes have the advantage of not having a fixed price encoded in the tag, such that a scanner at the checkout line is employed to identify the price of each product from a computer system. However, some strategies demand manpower and material resources, or impose limited customer choices for selecting, packaging and weighting their products.

Several approaches have been proposed in the literature for fruit and vegetable recognition based on color, texture and shape features individually. Some previous works based on only one feature have not met certain application requirements, such as recognition accuracy rates. More recently, feature fusion using texture, color and shape have been applied to fruit and vegetable recognition, achieving superior results [3, 4].

A novel and effective automatic fruit and vegetable recognition methodology, based on fusing CENTRIST features and a Hue-Saturation Histogram representation, is proposed and analyzed in this work. Initially, background subtraction is performed on the images. Then, a set of features is extracted and fused. Finally, fruit and vegetable species are recognized.

The remainder of the paper is organized as follows. Section 2 briefly describes the Census Transform Histogram (CENTRIST) and Color CENTRIST. Section 3 presents the methodology proposed in this work, describing the preprocessing, the feature extraction, the feature reduction and classification stage. Section 4 describes and discusses the experiments and results. Section 5 concludes the paper with final remarks.

2 Background

Concepts related to the topic under investigation are briefly described in this section.

2.1 CENTRIST

CENsus TRansform hISTogram (CENTRIST) [8] was originally proposed for recognizing topological places and scene categories. This visual descriptor provides an accurate and stable performance on various scene images. Based on Census Transform (CT), this operator is a holistic representation, which models distribution of local structures. It is characterized for providing a high generalization for categorization, suppressing detailed textural information.

The descriptor compares the intensity value of a central pixel with its eight neighboring pixels. If the central pixel intensity value is higher than or equal to one of its neighboring pixels, bit 1 is set to the corresponding location, otherwise bit 0 is set. The resulting 8 bits can be concatenated in any order, converting the 8-bit binary representation to a base-10 number, being the Census Transform value (CT value) of the current central pixel. After computing a CT value for each pixel, a histogram descriptor of CT values is constructed with 256 bins, called CENTRIST.

2.2 Color CENTRIST

The Color CENTRIST [2] was introduced to embed color information into CENTRIST framework. It has proven to improve recognition accuracy rates for color scene categorization. This operator describes color information by working over

the hue-saturation-value (HSV) color space, where the three channels are normalized into the range of 0 to 255, requiring 8 bits to represent each channel. In order to propose a comparable representation to CT values of CENTRIST, the Color CENTRIST devises a framework for representing color information by 8 bits and different quantization for the three channels.

To represent hue, saturation and value components, 1, 2 and 5 bits are allocated, respectively. The three channels are uniformly quantized. The hue axis is divided into $2^1 = 2$ ranges, i.e., $[0, 127]$ and $[128, 255]$. The saturation axis is divided into $2^2 = 4$ ranges, i.e., $[0, 63]$, $[64, 127]$, $[128, 191]$, $[192, 255]$. The value axis is divided into $2^5 = 32$ ranges, i.e., $[0, 7]$, $[8, 15]$, ..., $[240, 247]$, and $[248, 255]$. The hue component for a pixel is transformed into a color index (base 10) by $i_h = [\frac{h \times 2^1}{256}]$, then the resulting value is converted into a binary number through 1 bit. In a similar way, the color indices for saturation and value components are represented by $i_s = [\frac{s \times 2^2}{256}]$ and $i_v = [\frac{v \times 2^5}{256}]$, then transformed into binary numbers through 2 and 5 bits, respectively.

This operator gives the highest priority to value component. Hence, the color indices (base 2) of a pixel are concatenated as $(i_v i_s i_h)$. Then, the binary stream is converted into a base-10 system. Thereby, the CENTRIST operator is applied over the resulting color indices (base 10), which generates a histogram descriptor of CT values of length 256.

3   Methodology

The proposed methodology for fruit and vegetable recognition is composed of four main stages: (i) preprocessing, (ii) feature extraction, (iii) feature reduction, and (iv) classification. These steps are described as follows.

3.1   Preprocessing

The image preprocessing stage is crucial for the recognition task, whose main purpose is to perform background subtraction [1]. Background subtraction is basically used for segmenting out objects of interest in a scene. This preprocessing procedure consists of the following six steps: (i) conversion of input image into the HSV color space [7]; (ii) performing Sobel operator over saturation component; (iii) filling small holes using closing morphological operator with a disk structuring element; (iv) removing noise using median filtering; (v) finding a rectangular mask to find an interest region; and (vi) cropping the interest region. Figure 1 illustrates the stages in some samples of images.

3.2   Feature Extraction

Three types of features – CENTRIST [8], Color CENTRIST [2] and Hue-Saturation Histogram – were extracted and some of them fused for fruit and vegetable recognition.
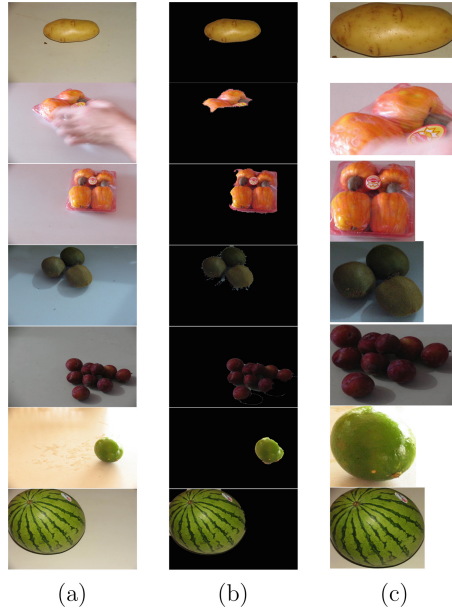
Fig. 1. (a) Original images from the Supermarket Produce data set [6]; (b) interest region detection from (a); (c) cropped interest regions from (b).

The CENTRIST [8] features were extracted from the entire images, forming a vector of length 256. This feature vector was normalized by dividing each intensity level r by the total pixels n of the image, as follows

$$p(r) = \frac{h(r)}{n} \tag{1}$$

where $h(r)$ defines the occurrence frequency of each intensity level in the image, and $p(r)$ gives the probability of occurrence intensity. CENTRIST is able to capture local structures of the image. Figure 2 shows a sample of a census transformed fruit image.
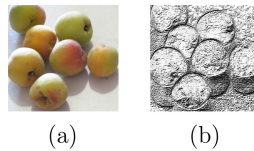
Fig. 2. (a) Cropped image from the supermarket produce data set [6]; (b) census transformed image from (a).

In a similar way to CENTRIST feature extraction, we applied the Color CENTRIST [2] operator over the entire image. The resulting feature vector was normalized, forming a vector of 256. Figure 3 illustrates the process.
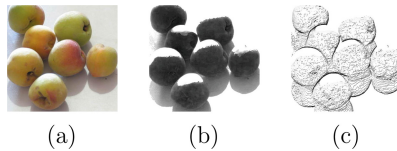
Fig. 3. (a) Cropped image from the supermarket produce data set; (b) HSV channels from (a) transformed into color indices; (c) census transformed image from (b).

After transforming all images from the RGB to the HSV color space, we performed several experiments over the different HSV channels and decided to generate the Hue-Saturation Histogram. From both hue and saturation channels, we computed two histograms, resulting in a feature vector of length 512 ($= 2 \times 256$). We also performed histogram normalization (Eq. 1).

3.3   Feature Reduction and Classification

Two techniques for feature reduction, Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), were sequentially used. Thereby, we applied this procedure individually for each feature set. We then used K-Nearest Neighbors (K-NN) and Support Vector Machine (SVM) classifiers for comparing the recognition rates.

4   Experimental Results

The proposed methodology was tested on the available Supermarket Produce data set proposed in [6]. This data set contains 2633 images of 15 different fruit and vegetable categories: Agata Potato (201), Asterix Potato (182), Cashew (210), Diamond Peach (211), Fuji Apple (212), Granny-Smith Apple (155), Honeydew Melon (145), Kiwi (171), Nectarine (247), Onion (75), Plum (264), Taiti Lime (106), Watermelon (192), and Williams Pear (159).

The Supermarket Produce data set comprises illumination deviations (Fig. 4), pose differences, variability in the number of elements and different maturity degrees (Fig. 5). Furthermore, this data set contains elements packed into plastic bags, creating reflections and shadows, as well as other occlusions, such as hands and cropping elements (Fig. 6).

Initially, we randomly choose 80% of samples of each class for the training set and the remaining 20% for the testing set. We set 20 different randomized data collections to conduct experiments. Thus, we performed experiments using CENTRIST, Color CENTRIST, HS-Histogram, and the fusion of CENTRIST

Fig. 4. Different illumination deviations for Asterix Potato category.

Fig. 5. Distinct maturity degrees for Williams Pear category.

(a)   (b)   (c)
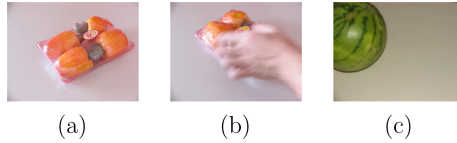
Fig. 6. Examples of occlusions. (a) cashew's plastic repository; (b) hand presence; (c) cropping Watermelon element.

with HS-Histogram, through six methods: K-NN, PCA+K-NN, PCA+LDA+K-NN, SVM, PCA+SVM and PCA+LDA+SVM. The results are shown in Table 1, whose values represent the fruit and vegetable recognition accuracy rates by running 20 experiments with different data collections.

From our experiments, we can notice that the PCA+LDA approach achieves better recognition accuracy than just using PCA. Moreover, we can observe that the fruit and vegetable recognition using the fusion of CENTRIST and HS-Histogram representation is much higher when compared to the other texture and color representations, especially when is followed by SVM classifier. From Table 1, it is possible to see that the CENTRIST representation works much

Table 1. Average accuracy, in percentage, using CENTRIST, Color CENTRIST, HS-Histogram, and the fusion of CENTRIST with HS-Histogram, for the Supermarket Produce [6] data set.

Recognition method	CENTRIST (%)	Color CENTRIST (%)	HS-Histogram (%)	CENTRIST+ HS-Histogram(%)
K-NN	79.58	69.90	88.12	92.40
PCA + K-NN	79.58	71.09	89.51	92.56
PCA + LDA + K-NN	85.96	76.04	90.50	96.60
SVM	86.18	74.85	90.50	**97.23**
PCA + SVM	89.15	78.02	91.29	94.47
PCA + LDA + SVM	**90.64**	**80.20**	**91.89**	95.96

better than the Color CENTRIST representation, although the latter considers further information, as color features. We can see that the accuracy using HS-Histogram is slightly better than using a CENTRIST representation.

Table 2 shows the confusion matrix using the fusion of CENTRIST and HS-Histogram features for SVM classifier. It is possible to observe that accuracy rates up to 90% are achieved for each class. Furthermore, the results show that the proposed approach is able to proper recognize the variety of fruit and vegetable species, for instance, Fuji Apple and Granny-Smith Apple, and Diamond Peach and Nectarine.

Table 2. Confusion matrix using fusion of CENTRIST and HS-Histogram representation with SVM classifier for the Supermarket Produce data set.

	Agata Potato	Asterix Potato	Cashew	Diamond Peach	Fuji Apple	Granny-Smith Apple	Honeydew Melon	Kiwi	Nectarine	Onion	Orange	Plum	Taiti Lime	Watermelon	Williams Pear	Sensitivity (%)
Agata Potato	40	0	0	0	0	0	0	0	0	1	0	0	0	0	0	97.56
Asterix Potato	0	37	0	1	0	0	0	0	0	0	0	0	0	0	0	100.00
Cashew	0	0	42	0	1	0	0	0	0	0	0	0	0	0	0	97.67
Diamond Peach	0	0	0	40	0	0	0	0	1	0	0	0	2	0	0	93.02
Fuji Apple	0	0	0	1	42	0	0	0	0	0	0	0	0	0	0	97.67
Granny-Smith Apple	0	0	0	0	0	32	0	0	0	0	0	0	0	0	0	100.00
Honewdew Melon	0	0	0	0	0	0	29	0	0	0	0	0	1	0	0	96.67
Kiwi	3	0	0	0	0	1	0	32	0	0	0	0	0	0	0	91.43
Nectarine	0	0	1	0	0	0	0	0	49	0	0	0	0	0	0	98.00
Onion	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	100.00
Orange	0	0	0	0	0	0	1	0	0	0	19	0	0	1	0	90.48
Plum	0	0	0	0	0	0	0	0	0	0	0	53	0	0	0	100.00
Taiti Lime	0	0	0	0	0	0	1	0	0	0	0	0	31	0	0	96.88
Watermelon	0	0	0	0	0	0	0	0	0	0	0	0	0	22	0	100.00
Williams Pear	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	100.00
Precision (%)	93.02	100.00	97.67	97.56	97.67	100.00	93.55	100.00	98.00	94.12	100.00	100.00	91.18	95.65	100.00	**97.23**

Table 3. Comparison of average accuracy rates (%) for fruit and vegetable recognition.

Approach	Strategy	Accuracy (%)
Proposed	CENTRIST and HS-Histogram representation + SVM	**97.23**
Rocha et al. [6]	Bagging Ensemble of Linear Discriminant Analysis	97.00
Dubey et al. [3]	Global Color Histogram and Local Binary Patterns + SVM	93.84
Arivazhagan et al. [1]	Minimum distance criterion	86.00

We compared our proposed method to others available in the literature. Table 3 summarizes the comparison results on the Supermarket Produce data set. We can see that the proposed approach, i.e., the fusion of CENTRIST and HS-Histogram representation with SVM, obtains superior results. Table 3 is sorted in descending order by accuracy rate.

5   Conclusions

In this work, experimental results have shown that the fusion of CENTRIST features with HS-Histogram can achieve high accuracy rates for fruit and vegetable recognition. Our method has also demonstrated to be robust due the presence of different occlusions, for instance, shadows, plastic bags, cashier's hand.

Despite the fact that the PCA and LDA approach did not provide a higher accuracy rate when combining CENTRIST and HS-Histogram, this approach allowed to reduce and select discriminative features, as well as increasing recognition accuracy significantly when features were used individually. Furthermore, computing CENTRIST and HS-Histogram features does not demand high computational resources. The generation of the combined feature vector is a very fast and efficient process, allowing fruit and vegetable recognition in real time.

References

1. Arivazhagan, S., Shebiah, R.N., Nidhyanandhan, S.S., Ganesan, L.: Fruit recognition using color and texture features. J. Emerg. Trends Comput. Inf. Sci. **1**(2), 90–94 (2010)
2. Chu, W.T., Chen, C.H., Hsu, H.N.: Color CENTRIST: embedding color information in scene categorization. J. Vis. Commun. Image Representation **25**(5), 840–854 (2014)
3. Dubey, S.R., Jalal, A.S.: Fruit and vegetable recognition by fusing colour and texture features of the image using machine learning. Int. J. Appl. Pattern Recogn. **2**(2), 160–181 (2015)
4. Gao, G., Zhao, S., Zhang, C., Yu, X., Li, Z.: Study on fruit recognition methods based on compressed sensing. J. Comput. Theor. Nanosci. **12**(9), 2937–2942 (2015)
5. Leemans, V., Destain, M.F.: A real-time grading method of apples based on features extracted from defects. J. Food Eng. **61**(1), 83–89 (2004)
6. Rocha, A., Hauagge, D.C., Wainer, J., Goldenstein, S.: Automatic fruit and vegetable classification from images. Comput. Electron. Agric. **70**(1), 96–104 (2010)
7. Sural, S., Qian, G., Pramanik, S.: Segmentation and histogram generation using the hsv color space for image retrieval. In: International Conference on Image Processing, vol. 2, pp. 589–592. IEEE (2002)
8. Wu, J., Rehg, J.M.: CENTRIST: a visual descriptor for scene categorization. IEEE Trans. Pattern Anal. Mach. Intell. **33**(8), 1489–1501 (2011)
9. Zawbaa, H.M., Abbass, M., Hazman, M., Hassenian, A.E.: Automatic fruit image recognition system based on shape and color features. In: Hassanien, A.E., Tolba, M.F., Taher Azar, A. (eds.) AMLTA 2014. CCIS, vol. 488, pp. 278–290. Springer, Heidelberg (2014). doi:10.1007/978-3-319-13461-1_27