Published in final edited form as:
SenseCam 2013. 2013 November ; 2013: 80–81. doi:10.1145/2526667.2526681.

Eating Activity Detection from Images Acquired by a Wearable Camera

Xin Sun,
Harbin Institute of Technology

Hongxun Yao,
Harbin Institute of Technology

Wenyan Jia, and
University of Pittsburgh

Mingui Sun
University of Pittsburgh

Abstract

We present an eating activity detection method via automatic detecting dining plates from images acquired chronically by a wearable camera. Convex edge segments and their combinations within each input image are modeled with respect to probabilities of belonging to candidate ellipses. Then, a dining plate is determined according to a confidence score. Finally, the presence/absence of an eating event in an image sequence is determined by analyzing successive frames. Our experimental results verified the effectiveness of this method.

Keywords

Activity detection; dining plate detection; eating event detection; wearable computer

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

As the epidemic of obesity spreads rapidly over the world, the study of people's diet becomes extremely important. We have developed a wearable computer, eButton [3], to record human activities, including diet. This device chronically records multimodal data in front of the wearer. Since the dataset recorded by eButton is extremely large, it is highly desirable to detect eating activity automatically. Though some food recognition methods have been developed [4], direct recognition of a large variety of human food is very difficult. We thus turn to automatic detection of elliptic patterns representing circular dining plates. It is clear that this approach does not detect all eating events. In cases where a plate is not used during eating, a manual data processing method can be used.

Numerous ellipse detection algorithms have been developed. The least-square methods [2] are popular choice for ellipse fitting due to their linearity and simplicity. Ellipse detection based on RANdom SAmple Consensus (RANSAC) [1] divides an image content into small subsets and takes samples of edges from each subset to create the most probable elliptic parameter set. Then, a cost function is used to evaluate the presence of an ellipse by counting the number of points in the proximity of the parameter set.

In this paper, we first localize the elliptic pattern representing a plate. Specifically, we extract edges of the input image and split these edges into segments according to a convexity criterion. By evaluating these segments, we select the ones with large probabilities of belonging to a plate. Then, different combinations of the selected edge segments are compared in order to reconstruct an ellipse representing the plate. Finally, the occurrence of an eating event is detected by analyzing a sequence of successively recorded image frames.

METHODS

Given an input image I, we extract its edge information by computing the gradient:

$$E(x, y) = |\nabla[G_\sigma(x, y) * I(x, y)]|^2 \quad (1)$$

where ∇ denotes the spatial gradient operator, * denotes convolution and G_σ is a Gaussian filter with standard deviation σ. We split the edge curve into segments according to changes in monotonicity of the edge curve in either x-direction or y- direction. Let $Q = \{q_i\}_{i=1}^{|Q|}$ be the set of edge segments. For each segment q_i, we find the best fitting ellipse U_i using the least-square algorithm [2]. Then, the probability of edge segment q_i belonging to an ellipse is modeled as:

$$p(q_i \mid U_i) \propto F_c(q_i, U_i) \quad (2)$$

where F_c is the number of points in I where q_i intersects with elliptic curve U_i, and ∝ provides a necessary normalization so that $p(q_i|U_i)$ represents a probability. Based on probabilities (2), the top-N edge segments with the largest probabilities are selected from Q, i.e. $\hat{Q} = \{q_j\}_{j=1}^{N}$. Hance, $\hat{Q}$ contains all major edges of the plate. Let D denote the set of edges belonging to a plate. Normally, we have $D \subset \hat{Q}$. In order to find D, we evaluate each combinations of edges in $\hat{Q}$. For each combination $C_k = \{q_{((c))}\}_{c=1}^{|C_k|}$, we compute the best fitting ellipse U_{C_k} based on the combined edges in C_k. Then, the degree of consistency between the fitting ellipse and combined edges in C_k is modeled by following probability

$$p(C_k \mid U_{C_k}) \propto \sum_{q \in C_k} F_c(q, U_{C_k}) \quad (3)$$

Since the expected combination of edge segments belonging to a plate should have the largest portion of intersections with the corresponding fitting ellipse, we select the combination with the largest probability given by (3) as the resulting plate edge set D, and treat this value as a confidence score of the estimated plate location. Fig.1 illustrates the concepts of our plate location procedure.

This procedure always finds a plate candidate regardless whether image I contains a plate. In order to judge the existence of a plate in I, we set a thresholds T_p (determined empirically) to evaluate the confidence score defined in (3). If this score is larger than T_p, we determine I contains a plate.

For an image sequence acquired by a wearable camera, we apply the above plate detection procedure to each frame in the sequence. If the result of plate detection is positive in a number of adjacent frames, we regard that an eating event is detected. Note that this number, which can be optimized experimentally, is related to the rate of image acquisition by the wearable camera.

EXPERIMENTAL RESULTS

Dataset:

We used a self-constructed wearable computer called eButton to collect data. This multi-sensor device was designed to evaluate both diet and physical activity objectively [3]. In our experiment, the camera (in 120° viewing angle) of the eButton was set to collect images in the resolution of 640×480 pixels at a rate of every 2-4 seconds. With an institutional review board (IRB) approval, five human subjects participated in the experiment in which their daily lives in a two-week period were recorded. The total number of images obtained in this experiment was over 300,000.

Performance:

We implemented the plate detection method on the wearable camera captured images. Table 1 summarizes our results. Table 1(a) shows the accuracy in plate detection from all individual images where TP, FN, FP and TN denote, respectively, true positive, false negative, false positive, and true negative cases. The values have been normalized with respect to the number of plate-containing images (for TP and FN) or the number of non-plate images (for FP and TN). Table 1(b) shows the eating detection accuracy based on successive frames in the eButton acquired image sequences. Again, each value has been normalized by the number of meals which was determined by visual inspection of the input sequences as the ground truth. Our results indicate that the errors in both plate and eating event detections are relatively small. Further inspections on the images producing erroneous results indicate that the undetected meals were mostly breakfasts which were either too short or not served using plates.

In order to exemplify our results, we show several challenging cases of plate detection in Fig. 2 where the plates were detected successfully despite the difficulties. Fig. 3 shows an example of a one-day data where plates were detected repeatedly in three periods (Fig. 3(a)) which indicate breakfast, lunch and dinner. Typical detection results in these three periods are shown in Fig. 3(b)-3(d).

CONCLUSION

We have presented an automatic eating activity detection method via analyzing elliptic patterns representing dining plates in wearable camera acquired image sequences. Edge segments belonging to an ellipse are modeled in terms of probabilities. Combinations of different edge segments are optimized by maximizing the fittings between edge segments and modeled ellipses. Eating events are detected by analyzing adjacent frames in an image sequence. Our experimental results have verified the effectiveness of the proposed method.

ACKNOWLEDGEMENT

This work was supported in part by the National Institutes of Health grants U01 HL91736 and R01CA165255.

REFERENCES

1. Bolles R, and Fischler M A ransac-based approach to model fitting and its application to finding cylinders in range data. In IJCAI (1981), 637–643.
2. Fitzgibbon A, Pilu M, and Fisher R Direct least square fitting of ellipses. TPAMI21, 5 (1999), 476–180.
3. Sun M, Fernstrom J, Jia W, and et al. A wearable electronic system for objective dietary assessment. In American Dietetic Association (2010), 45–17.
4. Yang S, Chen M, Pomerleau D, and Sukthankar R Food recognition using statistics of pairwise local features. In IEEE Conf on CVPR (2010), 2249–2256.

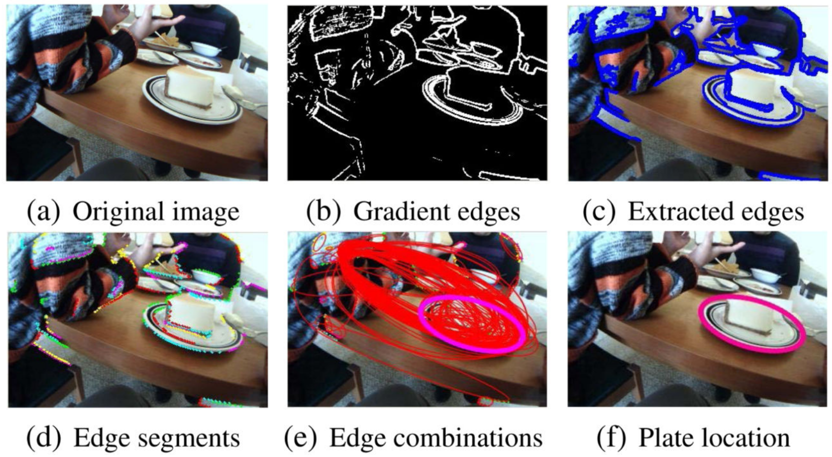


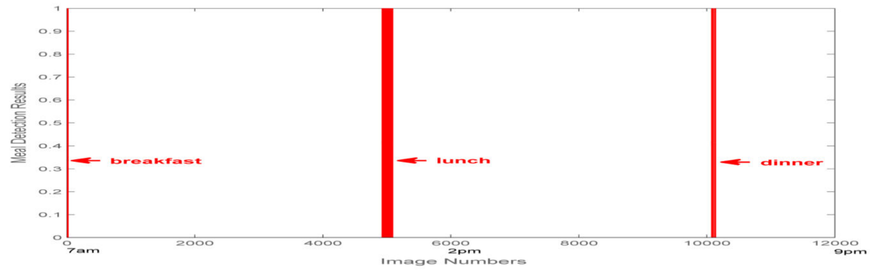
Figure 1.
Dining plate location procedure.

(a) Partial occlusion (b) Fast moving (c) Presence of utensils

Figure 2.

Plate localization in several challenging images

Author Manuscript
Author Manuscript
Author Manuscript



(a) Plate detection outputs of one-day data

(b) Breakfast (c) Lunch (d) Dinner

Figure 3.

Typical one-day plate detection results in three meals

Table 1.

Eating activities detection results on two-week personal data.

TP	FN	FP	FN	Turly Detected	Undetected	False detected
0.9118	0.0882	0.0530	0.9470	0.8947	0.1053	0.0526
(a) Plate detection accuracy				(b) Eating event detection accuracy		