Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Food Engineering 78 (2007) 246–256
JOURNAL OF FOOD ENGINEERING
www.elsevier.com/locate/jfoodeng

A computer vision system for appearance-based descriptive sensory evaluation of meals

Per Munkevik [a,b], Gunnar Hall [a], Tom Duckett [b,*]

[a] Swedish Institute for Food and Biotechnology, Box 5401, SE-402 29 Göteborg, Sweden
[b] Centre for Applied Autonomous Sensor Systems, Department of Technology, Örebro University, SE-70182 Örebro, Sweden

Received 5 March 2005; accepted 26 September 2005
Available online 20 December 2005

Abstract

This paper presents a complete machine vision system for automatic descriptive sensory evaluation of meals. A human sensory panel first developed a set of 72 sensory attributes describing the appearance of a prototypical meal, and then evaluated the intensities of those attributes on a data set of 58 images of example meals. This data was then used both to train and validate the performance of the artificial system. This system covers all stages of image analysis from pre-processing to pattern recognition, including novel techniques for enhancing the segmentation of meal components and extracting image features that mimic the attributes developed by the panel. Artificial neural networks were used to learn the mapping from image features to attribute intensity values. The results showed that the new system was extremely good in learning and reproducing the opinion of the human sensory experts, achieving almost the same performance as the panel members themselves.

Keywords: Artificial neural network; Sensory descriptive evaluation; Bayes classifier; Appearance; Meal; Image analysis

1. Introduction

Industrially prepared ready meals are a growing market in Europe. One of the challenges for the industry is to make the production more automatic and flexible, ideally by using autonomous robotic systems. However, as yet, there is no automatic system to identify, classify or grade the components of a meal or a complete meal in the same way as a human being, using his/her senses.

Human sensory evaluation is widely used in the food sector to measure and understand the responses given by the human senses. There are different types of evaluations depending on the goal of the evaluation; product quality control, consumer acceptance or to obtain the intensities of the attributes describing the perception of a product. To measure the intensities of product specific attributes, descriptive sensory evaluation is used. The resulting intensities from such an evaluation are used to compare two or more products or to compare the product against a specification that the product should match (Lawless & Heymann, 1999).

In order to automate, sensor systems are needed that can be used within the production of food products. This study will introduce an automatic method for descriptive sensory evaluation of a complex food product (a meal) using a computer vision system, comprising algorithms for image analysis, including image segmentation, feature extraction and pattern recognition. The system uses training data acquired from a human sensory panel in order to reproduce the panel's intensity judgements. The result is a complete artificial vision system that automatically gives the sensory intensities from digitised colour images of meals, and which closely follows the performance of the human sensory panel. Our experiments showed that the performance of the new system was comparable to that

* Corresponding author. Tel.: +46 19 303000; fax: +46 19 303463.
E-mail address: tom.duckett@tech.oru.se (T. Duckett).

0260-8774/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jfoodeng.2005.09.033

of the sensory panel, performing only slightly worse than the least consistent of the human panel members. We calculated the percentage of intensity levels that lay within the 95% confidence interval of the whole panel opinion, and found that the artificial system obtained 82% compared to 86% for the least consistent human panel member.

The rest of this paper is structured as follows. After a brief review of related work, the acquisition of images and sensory intensity judgements from a human sensory panel is described in Section 2. Section 3 describes the complete vision system, including a novel method for improving image segmentation (Section 3.1) and a novel method for extracting features corresponding to the sensory attributes selected by the panel (Section 3.3). In the last two chapters the results are presented followed by conclusions and some thoughts on possible future developments.

1.1. Related work

Several studies have been done on sensor systems as a complement to or as replacement of the human(s) in the evaluation of food products. In these studies electronic noses, electronic tongues and vision systems were used. There are cases when one type of sensor is used, as in the study presented by Park, Kim, and Noh (2002), where an electronic nose was used to classify the quality of stored soymilk. Others use sensor fusion, i.e., two or more types of sensors in combination, as in the work done by Korel, Luzuriaga, and Balaban (2001), where an electronic nose and a vision system was used for the quality assessment of fish.

In much work the goal is quality control of a specific product or of a step in the production of a food product, e.g., Sun and Brosnan (2003). However, only a few studies examine the performance of artificial sensor systems for descriptive sensory analysis of products. Interesting examples are the work done by Bleibaum et al. (2002), where they compare the results from a descriptive sensory panel with measurements from an electronic nose and an electronic tongue, and the study performed by Gao, Tan, Shatadal, and Heymann (1999), where they looked into the possibilities to predict the sensory properties of extruded puff texture by digital image analysis.

2. Data acquisition

2.1. Image acquisition

The components used in the meal in this study were potatoes, meatballs, sauce, jam and vegetables, and they were mounted by an experienced meal designer on white plates (diameter 260 mm). During mounting some aspects of the meal were varied by the designer; how and where each component was placed, the amount of the potato component and the amount of the meatball component. Two example meals can be seen in Fig. 1.

The meals were each placed on a red tray to obtain a uniform background and photographed using an USB PC-camera (Philips TOUCAM PCVC740K) and the software provided with the camera. The camera was positioned 0.5 m above the plate to get images containing all the components on the plate with as good resolution as possible. The process resulted in 58 images of meals with the same components with differences in the previously mentioned aspects. For some meals the plates were rotated and for others the images were electronically mirrored with no variations in the otherwise considered aspects. The images were stored at a resolution of 1280×960 pixels as bitmap images with a pixel value range of 0–255 (8-bit) and colour printed on white paper for the human sensory evaluation. The colour representation used for the images was RGB (red, green, blue).

2.2. Acquisition of training data from a human sensory panel

Quantitative sensory descriptions of the meal images were obtained using descriptive sensory analysis (Lawless & Heymann, 1999). In this study a panel with five experienced sensory judges was used. The judges had prior

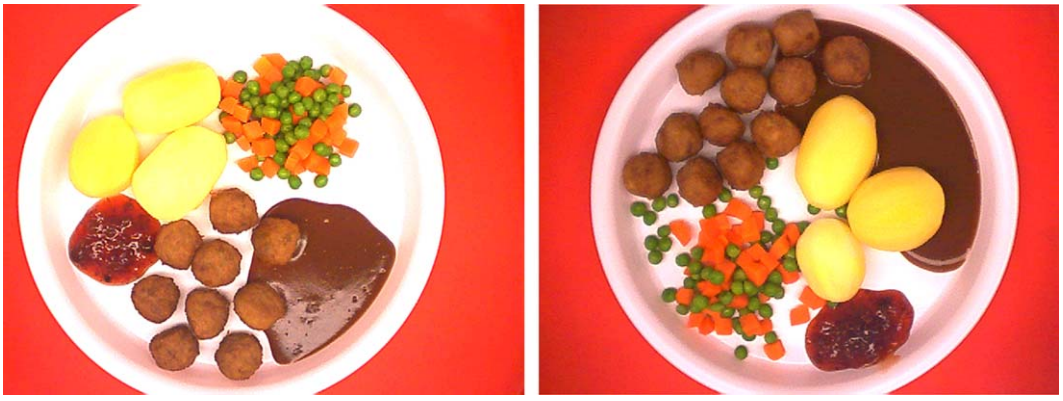

Fig. 1. Two samples of the meal images used showing some of the variations considered during mounting, e.g., the right meal contains larger amount of meatballs compared to the left one.