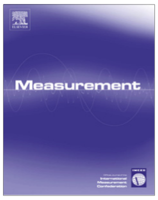
Measurement 65 (2015) 11–18

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement



Using microwave metrology to count calories

Michael A. Lexa [a,*], Iwan Njoto Sandjaja [b], Robert J. Marks II [b], Randall Jean [b], Kirk Marquard [a], William Platt [a], Aghogho Obi [a], V. Bogdan Neculaes [a], Jack M. Webster [a]

[a] *GE Global Research, Niskayuna, NY 12309, USA*
[b] *Electrical and Computer Engineering, Baylor University, Waco, TX 76798, USA*

ARTICLE INFO

Article history:
Received 15 September 2014
Received in revised form 26 December 2014
Accepted 30 December 2014
Available online 9 January 2015

Keywords:
Microwave spectroscopy
Calories
Nearest neighbor
Homogeneous mixtures
Moisture estimation
Food constituents

ABSTRACT

One non-destructive way to probe the caloric content of food is to examine the transmission spectra of low energy microwaves over a broad band (approximately 1–8 GHz). This paper proposes a simple, but generalizable, nearest neighbor scheme to estimate the calories of homogeneous mixtures of oil, sugar, and water. The approach's performance is empirically quantified and is also compared to the performance of an oracle estimator. We report that for this study the estimator achieves an average absolute error of approximately 10%. A heuristic extension of the nearest neighbor estimator is also discussed that can in some cases improve the average error performance. The work represents the first steps toward accurate and reliable calorie estimation techniques for complex, non-homogeneous foods of varying shapes and amounts.

1. Introduction

In this paper, we explore the problem of estimating the total caloric content of simple water-based mixtures using microwave spectroscopy. The overall approach is based on a simple, yet profound observation that the calories within most common foods can be well approximated by only knowing the item's mass and the fractions of water and fat comprising it. The observation, reported by Webster and Neculaes [1] (see also Neculaes et al. [2]), comes from a simple linear regression of more than 6500 foods listed in the United States Drug Administration (USDA) National Nutrient Database [3]. Specifically, the total calories c in a food item can be approximated by

$$c \approx (8.89p_f + 3.79(1 - p_w - p_f))w, \qquad (1)$$

where p_f and p_w denote the fraction of fat and water within the item, w denotes its weight/mass in grams, and the numerical constants have units of calories per gram (C/g). This expression says that, with the exception of fat, all basic food constituents (complex carbohydrates, protein, and simple sugars) have essentially the same caloric density; fat, on the other hand, has over twice as many calories per gram.[1] Thus (1) reduces the problem of calorie estimation to the problem of estimating the water and fat content of a food item. (We assume the weight/mass of a food item can be readily measured.)

Here we propose and analyze a nearest neighbor estimator that estimates the constituent percentages of simple oil–water–sugar mixtures and thus, via (1), yields calorie estimates. The approach is based on creating a dictionary of representative microwave transmission spectra and defining an appropriate distance metric that maps a test spectrum to the dictionary element that best represents it (see Section 3 for details). For the class of mixtures studied, the estimator achieves average errors of 2.2%, 2.5%, 2.8% for oil, sugar, water respectively and 10.2% for calories. These results are on par with other estimators proposed in the literature for similar experiments [4–6] and are roughly

---

* Corresponding author.
E-mail address: lexa@ge.com (M.A. Lexa).

[1] We assume that water as a basic food constituent has no calories.

http://dx.doi.org/10.1016/j.measurement.2014.12.050

within one percentage point of the performance of an oracle estimator for the constituent estimates and three percentage points for calories. This indicates that while some improvement may be gained by considering different distance metrics than the one proposed here, more sophisticated estimation methods will be necessary to improve average errors below approximately 1.4% for constituents and 7% for calories. However, for applications that can tolerate these levels of errors, the nearest neighbor estimator offers conceptual and computational simplicity.

1.1. Motivation

With the increasing prevalence of obesity worldwide and the growing associated costs (estimated to be $147 billion USD in the US alone in 2008 [7]), a device that could reliably measure the calories in foods would presumably help individuals track caloric intake and help them maintain a healthy weight. Currently there are several commercial offerings able to monitor *burned* calories [8–10], but there are few, if any, precise and easy-to-use devices that measure caloric content of *non-homogeneous* foods (e.g. sandwiches and kebabs). Current optical devices (including near infrared) can yield accurate calorie counts of homogeneous foods [11], but because they only probe the food's surface, calorie estimation of non-homogeneous foods is problematic. Caloric intake can be tracked manually using publicly available nutritional information, but this approach is often tedious and error prone. The microwave spectroscopic approach presented here has the potential to overcome these problems and aims to fill this societal need.

1.2. Related work

In microwave spectroscopy, materials are delineated by differences in their complex permittivity. It is natural then to use the permittivity either directly or indirectly to infer information about a material's constituents. For example, Daschner et al. [6] estimate moisture content by performing a principal component regression over the measured dielectric spectra, i.e., over the real component of permittivity; Gibbs et al. [4] use time-domain features related to the received signal's delay, attenuation, and dispersion which are all characteristics related to complex permittivity; and Jean [12] leverages specific spectral characteristics to estimate the constituents of process materials. In contrast to these approaches which operate on a relatively small number of features ($\sim$10), the nearest neighbor estimator can be thought of as an estimator that uses a much larger set of features ($\sim$1000) because it operates on an entire measured spectrum. Alternatively, it can be thought of as an estimator only operating on a single object where the spectrum is thought of as being a signature of the food item and the goal is to search among a dictionary of signatures to find the best match.

2. Experimental data

Fig. 1 depicts the experimental setup. We used an Agilent vector network analyzer with two identical Cobham spiral antennas to measure the transmission response of the mixtures (the top antenna being the transmitter and the bottom being the receiver). Each antenna has an effective bandwidth of 1.8–18 GHz and each has low dispersion characteristics. The analyzer was calibrated from 0.9 GHz to its maximum frequency of 8.5 GHz, and all measurements were referenced to the response of an empty beaker. For each experiment the analyzer swept through a uniform set of frequencies and recorded the received spectrum.

The mixtures were contained in a cylindrical beaker with a height and diameter equaling 75 cm and 150 cm, respectively. In order to minimize the effects of multipath and ensure that most of the transmitted energy passed through the sample, we chose a beaker that has a diameter three times larger than the largest dimension of the spiral antenna. The beaker contained homogeneous mixtures of distilled water, soybean oil (Fisher Scientific catalog number S25622), and a sucrose solution. The sugar/sucrose was purchased from a grocery store and solubilized as a 50% by weight solution of sugar and distilled water. The resulting syrup when added to the mixtures was considered to be 50% water and 50% sugar.

In total, the responses of 64 mixtures were recorded. The first 49 were considered dictionary elements and were created by combining different concentrations of oil and sugar by weight in 5% increments up to a maximum concentration of 30%. Fig. 2 graphically illustrates the different percentages. The remaining 15 samples were mixtures whose constituent percentages were randomly chosen within the same range as the dictionary samples (see Table 1). These mixtures were used to quantify the performance of the proposed estimator. Both the temperature and the volume of the mixtures were kept constant.

The peaks in the spectra and their general shape are controlled in large part by the frequency response of the system antennas and by the geometry of the measurement setup (see Fig. 3 for three prototypical spectral responses). The changing complex electrical permittivity of the samples modulates the spectral shape according to the frequency dependent attenuation for the mixtures and by an effective lens action on the wave front that is produced by delay and dispersion as the wave propagates through the material under test and its glass beaker container. The soybean oil is a fairly low loss, non-dispersive propagation medium over the frequency range used for the experiment. The complex permittivity of pure water exhibits a Debye frequency response that represents an increase in attenuation or loss factor with increasing frequency, up

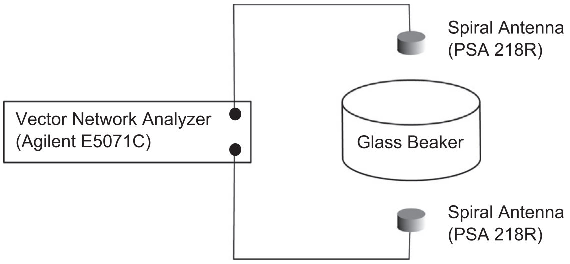


Fig. 1. Experimental setup.