34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Swallow Monitoring Through Apnea Detection in Breathing Signal*

Bo Dong, *Student Member, IEEE*, and Subir Biswas, *Senior Member, IEEE*

Abstract - **This paper presents the concepts, design, and algorithms for a wearable swallow monitoring system. Swallow monitoring can be used for assessing a person's overall food and drink intake habits as well as other swallow related disorders. The system works based on the key observation that a person's otherwise continuous breathing process is interrupted by a short apnea when she or he swallows as a part of food intake process. Using a wearable chest-belt, we detect swallows by detecting apneas extracted from breathing signal captured by the chest-belt. A matched filter based template matching framework is developed for swallow detection. A number of matched filter template design, both static and iterative, are developed for high-accuracy swallow detection.**

I. Introduction

In US, 68% of the population is considered overweight and/or obese [1], which is a condition that increases the risk of coronary heart disease, type-2 diabetes, and various types of cancers [2]. Two important components of obesity management are diet control and physical exercise. Traditionally, researchers have used self-reported questionnaires for estimating both food intake and physical activity levels for high-risk individuals. In recent years, accelerometer based instrumentation techniques [3] are starting to emerge as alternatives to self-reporting for physical monitoring. For food intake monitoring, however, not many instrumented efforts are reported in the literature. An instrumented system can eliminate subjectivity [4] associated with questionnaire based self-reporting systems.

Such an instrumented system can monitor the duration of each instance of food/drink intake, which can be used for estimating an individual's eating and drinking habits. Detecting a person's swallow events can enable such food/drink intake monitoring applications. An invasive method for swallow detection is Videofluoroscopy which is used to evaluate patients with neurological conditions affecting swallowing [5]. While providing ample information about different aspects of the swallow process, these methods are too involved and cumbersome to be used for everyday monitoring and food/drink intake analysis purposes.

Non-invasive methods use biological signals such as electromyography, sound, and movement to detect swallows. Surface electromyography (SEMG) and sound signal are used to detect the activation of muscles and the sound associated with swallow events [6]. The SEMG electrodes are normally attached to the bare skin in the neck region, which may raise user acceptability issues for prolonged usage due to cosmetic and safety reasons. A two-microphone system is developed in [7] for recording chewing and swallowing sound through the ear canal as well as externally through the air. Placing such microphones has similar cosmetic issues and therefore its suitability for prolonged usage is questionable. Respiratory Inductance Plethysmography (RIP) is used for swallow detection by measuring the airflow [8] in trachea. The RIP belts used for this method are often too involving to be useable for prolonged use in daily life settings.

In this paper we develop a wearable sensor system for swallow monitoring. The system works based on the key observation that during swallowing, because the trachea is blocked, a person is not able to breathe, causing a temporary *apnea*. Using a wearable chest-belt, we detect swallows by the way of detecting apneas extracted from breathing signal captured by the chest-belt. Since the belt can be worn inside, outside, or between garments (it does not need skin contact), it has the potential for prolonged comfortable daily usage without raising any cosmetic issues. After the swallow sequence is recorded, swallow pattern analysis can potentially be used for identifying non-intake swallows (or empty swallows), solid intake swallows, and drinking swallows.

II. Sensor Type and Characteristics

A piezo-respiratory belt contains a piezoelectric sensor placed between two elastic strips. Stretching the belt exerts a strain on the sensor, which generates a voltage proportional to the strength of the force. Comparing with other transduction principles, such as capacitive, inductive and resistive effects, piezoelectric phenomenon provides high sensitivity and excellent linearity over a wide amplitude range.

We use a piezo-respiratory chest-belt in our system for collecting breathing signal. Expansion and contraction of the chest during breathing creates a modulated voltage at the output of a belt-embedded piezo-electric sensor. This voltage represents the berating signal. The static response of the belt is shown in Fig. 1:a, which demonstrates good linearity of produced voltage as a function of belt elongation.

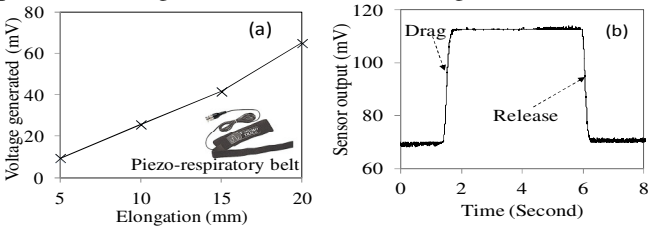

Fig. 1: Static and transient responses of piezo-respiratory belts

Fig. 1:b demonstrates the transient voltage signal captured by oscilloscope when the belt is first stretched by 15mm and then released after some time. It is notable that the output voltage closely follows the mechanical stimuli. The output signal is shaped (i.e. DC component removal, amplification) and then digitized using a 10-bit ADC with a sampling rate of 30Hz.

III. Breathing Signal and Apnea Analysis

Fig. 2 demonstrates two experimentally obtained breathing signal segments from one of the subjects. The ADC readings in the figure are directly proportional to the tension on the piezo-respiratory belt, meaning the rising edge in the graph

*This project is supported by an NIH grant R21 HL093395.

Bo Dong and Subir Biswas are with Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI, 48824, USA (e-mail:dongbo@egr.msu.edu, sbiswas@egr.msu.edu)

corresponds to inhalation and the falling edge corresponds to exhalation of a breathing cycle. As shown in the figure, a breathing cycle can be either normal (i.e. *Normal Breathing Cycle* or NBC) or elongated due to a momentary apnea caused due to a swallow event. A cycle that is elongated due to an apnea during the inhale process (see Fig. 2:a) is termed as *Breathing Cycle with Inhale Swallow* (BC-IS). Fig. 2:b shows swallows (i.e. apnea) at the beginning of an exhale which are termed as *Breathing Cycles with Exhale Swallow* (BC-ES).

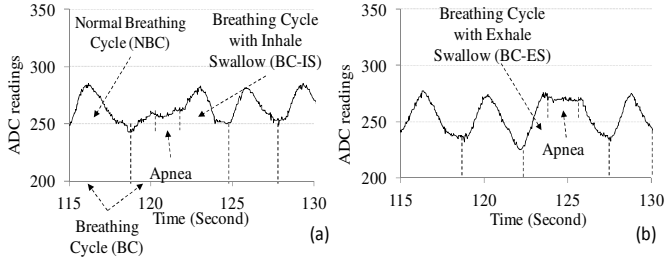

Fig. 2: Example of respiratory and swallow signals

IV. PROCESSING MECHANISM

Fig. 3 depicts the processing modules used for classifying breathing cycles towards swallow detection. The first step is to pass the ADC output through a low-pass filter for removing quantization noise caused by the A-to-D conversion process.

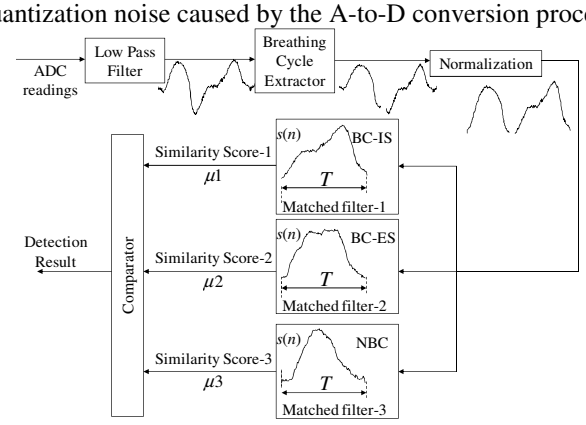

Fig. 3: Data processing architecture for swallow detection

The second step is to run the filtered data stream through a peak and valley detection module in order to extract the individual breathing cycles. The next processing module is used for normalizing the extracted cycles in both time and amplitude dimensions. The objective of normalization is to make sure that although different cycles may have different time and amplitude ranges (person-to-person and cycle-to-cycle), they can be effectively correlated with a single template. In other words, the normalization ensures that both $x(n)$ and $s(n)$ waveforms for the filter are of same duration.

The normalized breathing waveforms are then fed into three separate matched filters, each with a specific type of reference waveform. The filters use reference waveforms corresponding to *Normal Breathing Cycle* (NBC), *Breathing Cycle with Exhale Swallow* (BC-ES), and *Breathing Cycle with Inhale Swallow* (BC-IS), as shown in Fig. 3. The similarity score outputs of all three filters are compared to classify a breathing cycle as one of the above three types.

V. PERFORMANCE EVALUATION

Experiments with total seven subjects were carried out for this project. Representative results for three subjects are presented here..

A. Experimental Methods

Each subject performed three sessions, five minutes each. The subject was asked to wear the instrumented chest-belt and drink water with a swallow instruction given once in every 20 seconds. Each session resulted in approximately 80 Normal Breathing Cycle (NBC) and 15 breathing cycles with swallows (BC-ES and BC-IS). Data for all breathing cycles are wirelessly collected in a nearby notebook computer. A small microphone was also attached to the front part of the neck for recording the swallow sound. This audio signal, which was time-synchronized with swallow data extracted from the chest-belt, provided a control that was used for both training and verification of the proposed swallow detection mechanism. All subjects performed the above procedure.

The system performance is reported as ROC (Receiver Operating Characteristic) statistics of *True Positive Rate* vs. *False Positive Rate*. *True Positive Rate* is defined as the number of correctly detected BC-IS and BC-ES as a fraction of the total number of BC-IS and BC-ES. *False Positive Rate* is defined as the number of erroneously detected BC-IS and BC-ES as a fraction of the total number of NBCs.

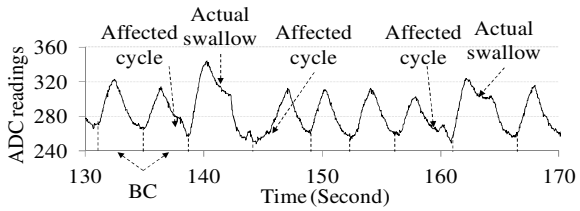

Fig. 4: Examples of BC modulation by adjacent swallows

B. Breathing Cycle Modulation by Adjacent Swallows

It was experimentally observed that sometimes a swallow event in a breathing cycle can modulate the signal for immediately adjacent cycles. Such modulations are demonstrated in the example trace in Fig. 4. It was determined that most of the affected cycles before actual swallows were caused by a minor change of breathing in subconscious anticipation of an impending swallow. The affected cycles after actual swallows were caused due to a similar reason. The subjects also reported that sometimes they executed a very minor second swallow for drinking any remaining liquid in the throat region. It turned out that the matched filter based system is often able to detect the above modulation and reported such occurrences as two swallows in consecutive breathing cycles, thus contributing to the false positive rate. Given that such modulations by adjacent swallows are always involuntary, it is reasonable to filter them out. Considering that it is rare for people to have real swallows in consecutive cycles, we counted two swallows in consecutive breathing cycles as a single swallow event.

VI. TEMPLATE WAVEFORM CONSTRUCTION

A. Arbitrarily Chosen Templates

We evaluated the detection performance with arbitrary combinations of NBC, BC-IS and BC-ES waveforms as the reference inputs in Fig. 3. For each subject, 3500 different *reference combinations* of NBC, BC-IS and BC-ES waveforms are created from approximately 300 collected breathing cycle waveforms. Then, for each *reference combination,* each of the 300 breathing cycle waveforms is classified to be one of three types using the system in Fig. 3.

Finally, by comparing these detection results with the actual breathing cycle types (observed from the neck-attached microphone) for all 300 cycles a *True Positive Rate* and a *False Positive Rate* is computed. An ROC pair is then computed for each of the 3500 *reference combinations*. Fig. 5 shows the ROC distribution of 3500 such pairs.

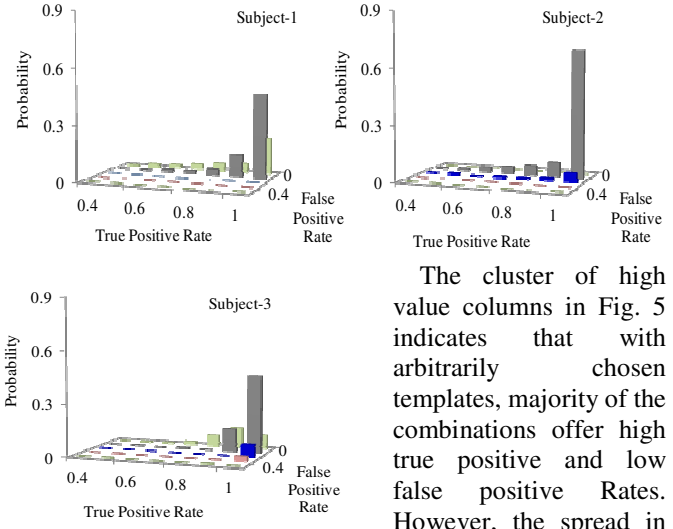

Fig. 5: ROC distribution with arbitrarily chosen templates

The cluster of high value columns in Fig. 5 indicates that with arbitrarily chosen templates, majority of the combinations offer high true positive and low false positive Rates. However, the spread in the distribution indicate that there exist NBC, BC-IS, and BC-ES waveforms which, if chosen as templates, can result in poor performance. Therefore, random template selection is not practical.

B. Existing Waveform Library

In this method, matched filter references are constructed by sample-by-sample averaging of waveforms available from a pre-constructed breathing waveform library. For a given subject, all normalized NBC waveforms from the library are averaged to construct the reference template for the NBC filter in Fig. 3. Similarly, all normalized BC-IS and BC-ES waveforms in the library are averaged for the creating the templates for the BC-IS and BC-ES filters respectively.

	True Positive Rate	False Positive Rate
Subject 1	0.933	0
Subject 2	0.933	0.089
Subject 3	0.911	0.010

Table 1: Performance with globally averaged templates

Table 1 summarizes detection performance with averaged template waveforms from an existing breathing waveform library, which is a prerequisite for this mechanism to work.

C. Controlled Breathing Cycles

Results in this section correspond to template waveforms that are computed based on the average of few known cycle types during a brief controlled phase. At the beginning of data collection, a subject is instructed to execute a fixed number of NBC, BC-IS, and BC-ES cycles. This provides the data for a set of few known cycles (i.e. *controlled cycle*) of each type. After the control phase, a sample-by-sample average NBC waveform is created from the recorded *controlled cycle*s of type NBC. This averaged waveform is then used as the template for the NBC matched filter in Fig. 3. Templates for BC-IS, and BC-ES filters are similarly constructed using the *controlled cycles* of the respective types. Fig. 6 reports detection performance using three *controlled cycles*. To capture the effects of variability present in the controlled cycles, for each cycle type we arbitrarily choose three breathing cycles from a library of cycles and use them as the *controlled cycles*. We compute the true positive and false positive rate for each such combination of three *controlled cycles* of all three cycle types. Performance distribution (i.e. ROC) of 3500 such combinations is reported in Fig. 6.

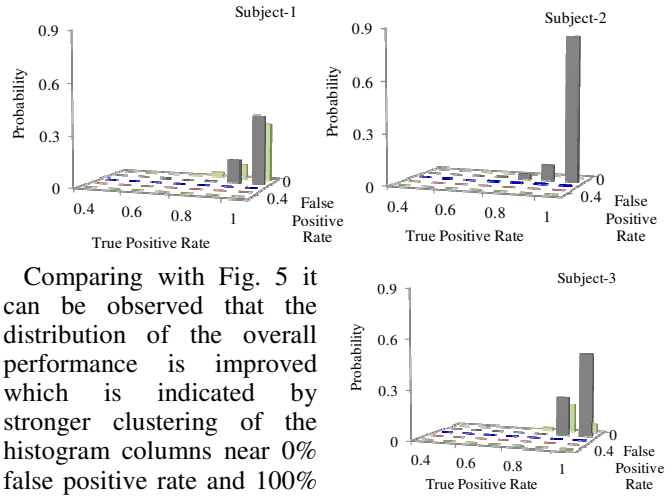

Fig. 6: ROC with templates using three controlled cycles

Comparing with Fig. 5 it can be observed that the distribution of the overall performance is improved which is indicated by stronger clustering of the histogram columns near 0% false positive rate and 100% true positive rate. It should be noted in Fig. 6 that although the performance distribution using the *controlled cycle* mechanism contain few very good points (i.e. the columns near 0% false positive rate and 100% true positive rate), there are also few low-performing columns in those distributions. Those indicate that bad quality *controlled cycles* can bring down the system performance.

Select initial templates T_{NBC}, $T_{BC\text{-}IS}$, and $T_{BC\text{-}ES}$
While (TPR and FPR are not stabilized)
  For (all collected breathing cycle x_i){
    Compute similarity scores μ_i^{NBC}, μ_i^{BC-ES}, μ_i^{BC-IS} for x_i:
    If $(\mu_i^{NBC} = \max(\mu_i^{NBC}, \mu_i^{BC-ES}, \mu_i^{BC-IS}))$
      x_i is NBC;
    If $(\mu_i^{BC-ES} = \max(\mu_i^{NBC}, \mu_i^{BC-ES}, \mu_i^{BC-IS}))$
      x_i is BC-ES;
    If $(\mu_i^{BC-IS} = \max(\mu_i^{NBC}, \mu_i^{BC-ES}, \mu_i^{BC-IS}))$
      x_i is BC-IS;
  }
  Generate a new set of templates as:
  T_{NBC} = average (detected NBCs with top 50% μ_i^{NBC})
  $T_{BC\text{-}ES}$ = average (detected BC-ESs with top 50% μ_i^{BC-ES})
  $T_{BC\text{-}IS}$ = average (detected BC-ISs with top 50% μ_i^{BC-IS})
End

Fig. 7: Algorithm for iterative template refinement

VII. ITERATIVE TEMPLATE REFINEMENT

In this section, we develop an iterative template refinement algorithm to address the above limitation. In this process, first, an NBC, a BC-IS, and a BC-ES waveform is chosen from the *controlled cycles.* Second, all collected breathing cycles are classified (see Fig. 3) using those three waveforms as the templates to the three matched filters. At this stage, each collected cycle is classified as NBC, BC-IS, or BC-ES. Third, all cycles that are classified as NBC are sorted based on their similarity scores obtained from the NBC matched filter in the

second step. Now, the top 50% of those NBC cycles are averaged to create the NBC template for the second iteration. The same process is also executed for BC-IS and BC-ES to form the templates for the second iteration. Seconds and third steps are then iteratively repeated till the true positive and false positive rates stabilize over consecutive iterations. The algorithm is summarized in Fig. 7.

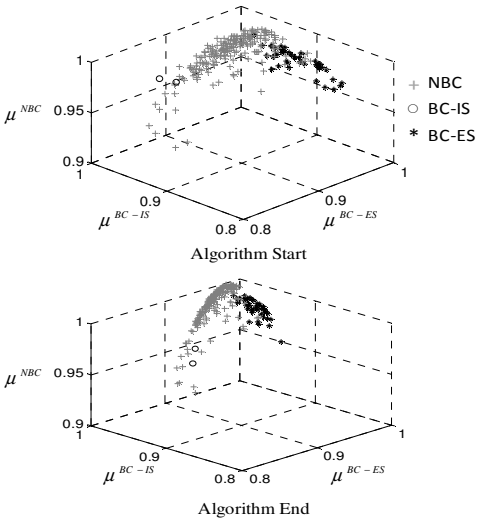

Fig. 8: Effects of iterative template refinement

The key concept here is that even when the initial template quality is not good, by choosing the top 50% of different types of the cycles, the algorithm should be able to iteratively refine the template quality, thus delivering good detection performance.

Fig. 8 depicts the algorithm dynamics in the form of the similarity score space at the start and end (i.e. at performance stabilization) of template refinement. The top graph shows the location of all the collected breathing cycles in the similarity score space obtained from the starting template waveforms.

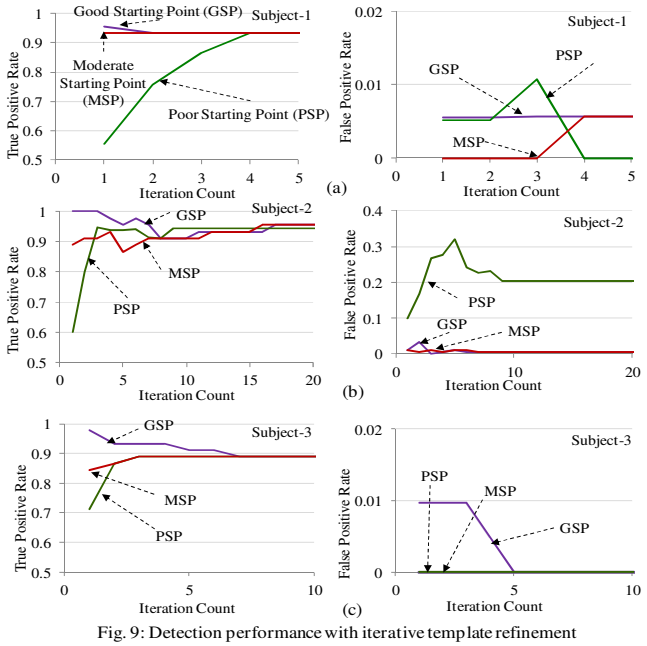

Fig. 9: Detection performance with iterative template refinement

The graph has 247 points, corresponding to 247 total collected breathing cycles. The bottom graph corresponds to similarity score space obtained from the template waveforms when the algorithm stabilizes. Observe that the overlapping among the three types of breathing cycles is much less in the bottom graph compared to the top one. This indicates a clear improvement of the template quality, leading to improved separation of different classified cycle types. The tighter clustering of the points in the lower graph provides additional indication to better template quality compared to the starting set. The patterns in Fig. 8 have been consistently observed for a wide range of initial template combinations.

Fig. 9 shows the performance of iterative template refinement based swallow detection for three different subjects. The evolution of true and false positive performance for each subject are reported for three different starting template sets, termed as, Good Starting Point (GSP), Moderate Starting Point (MSP), and Poor Starting Point (PSP). Observe that the true positive rate for PSP consistently improve with iterations for all three subjects. For MSP, such rates either improve or remain constant with iterations. With GSP, true positive rates go down slightly, although the decrement is always observed to be less than the improvements observed for PSP, thus establishing the effectiveness of this iterative template refinement approach. With few exceptions, the false positive rates generally go down with iterations. Closer investigation to those exceptions revealed that the reasons were highly deformed initial template sets that contained BC-ES or BC-IS waveforms representing affected adjacent swallows explained in Fig. 4.

VIII. Conclusion and Ongoing Work

This paper reported the concept, design, and algorithms for a wearable swallow monitoring system. A matched filter based template matching framework was developed. A number of template design mechanisms, both static and iterative, were developed for swallow detection with high true positive and low false positive performance. Ongoing work on this topic includes: a) validating detection mechanisms for both solid and liquid swallows, b) analyzing swallow sequences and durations for characterizing food and drink intake, and c) analyzing the artifacts caused by normal movement and behavior.

REFERENCES

[1] K. M. Flegal, M. D. Carroll, C. L. Ogden, and L. R. Curtin, "Prevalence and Trends in Obesity Among US Adults, 1999-2008," *JAMA: The Journal of the American Medical Association*, vol. 303, no. 3, pp. 235 –241, Jan. 2010.

[2] "Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults - NCBI Bookshelf." [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK2003/. [Accessed: 08-Feb-2012].

[3] Bo Dong and S. Biswas, "Wearable networked sensing for human mobility and activity analytics: A systems study," in *2012 Fourth International Conference on Communication Systems and Networks (COMSNETS)*, 2012, pp. 1–6.

[4] V. A. Vance, S. J. Woodruff, L. J. McCargar, J. Husted, and R. M. Hanning, "Self-reported dietary energy intake of normal weight, overweight and obese adolescents," *Public Health Nutr*, vol. 12, no. 2, pp. 222–227, Feb. 2009.

[5] J. Gates, G. G. Hartnell, and G. D. Gramigna, "Videofluoroscopy and Swallowing Studies for Neurologic Disease: A Primer1," *Radiographics*, vol. 26, no. 1, p. e22, Feb. 2006.

[6] O. Amft and G. Troster, "Methods for Detection and Classification of Normal Swallowing from Muscle Activation and Sound," in *Pervasive Health Conference and Workshops, 2006*, 2006, pp. 1–10.

[7] S. Passler and W.-J. Fischer, "Food Intake Activity Detection Using a Wearable Microphone System," in *2011 7th International Conference on Intelligent Environments (IE)*, 2011, pp. 298–301.

[8] A. Moreau–Gaudry, A. Sabil, G. Benchetrit, and A. Franco, "Use of Respiratory Inductance Plethysmography for the Detection of Swallowing in the Elderly," *Dysphagia*, vol. 20, no. 4, pp. 297–302, Mar. 2006.