WRLS-VFF-VT 알고리듬을 이용한 새로운 피치 검출 방법

이 교 식†·박 규 식††

요 약

본 논문은 WRLS-VFF-VT 알고리듬을 이용한 새로운 피치 검색 방법론을 제안하도록 한다. 제안된 알고리듬에서는 VFF(가변 망각 인자)를 사용하여 유성음에서의 주 여기 펄스 시점과 관련된 성문 폐쇄 점을 확인한다. 또한 본 논문은 VFF 기반 알고리듬과 함께 기존의 EGG 와 LP-Error 방법을 이용한 피치 검색 알고리듬에서 가변 한계 값을 이용 수정된 알고리듬을 제안한다. 제안된 알고리듬들은 주기와 주기 근간에서 강인한 피치 측정 능력과 준 주기 및 비 주기성 음성 신호에서도 우수한 피치 검색 기능을 기지고 있음을 알 수 있다. 제안된 알고리듬의 우수성을 입증하기 위해 실제 사람의 자연스러운 음성 및 사람의 비정상 상태 음성에서 준 주기 및 비 주기성 음성 진동 패턴을 확인하고 검출하는 성능 측정을 통하여 표준 SIFT 알고리듬과 비교 평가하였다.

A New Pitch Detection Method Using The WRLS-VFF-VT Algorithm

Kyo-Sik Lee† · Kyu-Sik Park††

ABSTRACT

In this paper, we present a new pitch determination method for speech analysis, namely VFF(Variable Forgetting Factor) based, by using the WRLS-VFF-VT(Weighted Recursive Least Square-Variable Forgetting Factor-Variable Threshold) algorithm. A proposed method uses VFF to identify the glottal closure points which correspond to the instants of the main excitation pulses for voiced speech. The modified EGG(ElectroGlottoGraphic) based, modified LP(Linear Prediction) Error based pitch determination algorithms with variable threshold are also presented in parallel with VFF based algorithm. We will show that these algorithms provide robust pitch estimates on a period by period basis and reliable pitch detection in quasi-periodic as well as in aperiodic speech signals. To verify the performance of the proposed pitch detectors, a standard pitch determination, a SIFT algorithm, is used to evaluate and compare the proposed algorithms for its capacity to detect and identify the patterns of aperiodic/quasi periodic voice vibration in natural human speech as well as in pathologic human speech.

1. Introduction

Pitch determination, viz. the detection and measurement of voice fundamental frequency(f_0) in natural human speech, still constitutes one of the most integral and at the same time most problematic areas of speech analysis. It has been established in numerous studies and even more so in practical

† 정 회 원 : 한세대학교 정보통신학과 교수
†† 정 회 원 : 상명대학교 컴퓨터정보통신학부 교수
논문접수 : 1998년 3월 12일, 심사완료 : 1998년 5월 29일

computer speech applications that the accurate representation of the voicing characteristics is of paramount importance for all aspects of speech signal processing. In speech coding, for instance, the quality of the vocoded speech deteriorates rapidly as a function of imprecise pitch estimates. In speech synthesis, considerable effort is dedicated today to the development and implementation of prosodic models for the generation of natural sounding pitch contours in text-to-speech.

Numerous studies have been dedicated to the design and evaluation of literally hundreds of pitch determination algorithms. Most of these studies have focused on pitch determination in stationary, quasi-periodic speech signals. Non-stationary, aperiodic signals, i.e. exhibiting occasional time and/or amplitude irregularities between successive periods of glottal excitation, have traditionally been disregarded and considered as reflecting pathological voice phenomena which are of no immediate concern to normal speech processing applications. However, aperiodic glottal vibrations occur far too often in normal, non pathological voices of both women and men, to be classified simply as a clinical syndrome or voice disorder. Aperiodic phonation has also been shown to be systematically employed by human speakers as an important demarcation cue in the continuous speech utterance. Standard pitch detection and pitch estimation algorithms usually fail to correctly identify patterns of aperiodic voice excitation, which are often wrongly classified either as voiceless stretches of speech, or associated with faulty pitch values, typically of the "octave error" type.

We can divide the speech based pitch determination algorithms(PDA) into two categories: time domain PDAs, such as the SIFT method, and short term PDAs, such as the autocorrelation and cepstrum methods[1][2]. The time domain PDAs are capable of providing accurate pitch periods, but are sensitive to signal degradations in the analysis frame. Short-term PDAs, on the other hand, are more robust, but provide only an average pitch estimate over a number of periods, because short term PDAs rely on the similarity of the speech signal over adjacent pitch periods. Thus, if a large number of pitch periods are contained in an analysis segment the estimated pitch value is a "smeared" or average value for the segment. Both methods lose all information about the absolute position of the glottal excitation.

The main objective of this paper is to develop a new pitch detection method, namely VFF based PDA using WRLS-VFF-VT(Weighted Recursive Least Square-Variable Forgetting Factor-Variable Threshold) algorithm. The modified EGG based and LP Error based PDAs with variable threshold are also proposed in parallel with VFF based algorithm[3][4]. We will show that these algorithms provide robust pitch estimates on a period by period basis and reliable pitch detection in quasi-periodic as well as in aperiodic speech signals.

This paper is organized as follows. In section II. we describes the WRLS-VFF-VT algorithm and implementation procedure. Section III. present modified EGG based, LP error based and the VFF based PDAs, but main emphasis on the VFF based algorithm. These methods are the time domain PDAs, and provide the pitch period on period by period basis. Section IV evaluate the performance of the proposed pitch detection algorithms. A standard time domain pitch determination, a SIFT algorithm, is used to evaluate and compare the proposed algorithms for its capacity to detect and identify the patterns of aperiodic/quasi-periodic voice vibration in natural human speech.

2. WRLS-VFF-VT Algorithm Description

A. Background for the WRLS Algorithm

We will assume that the speech signal is generated by an ARMA model represented by the following :

$$y_k = -\sum_{i=1}^{p} a_i(k) y_{k-i} + \sum_{j=1}^{q} b_j(k) u_{k-j} + u_k \qquad (1)$$

Where y_k denotes the k-th sample of the speech signal, u_k is the input excitation to the ARMA model, (p,q) are the order of the poles and zeros, respectively, of the ARMA model, and $a_i(k)$ and $b_j(k)$ are the time-varying AR and MA parameters, respectively. Model order selection techniques have been proposed by numerous authors [5][6][7]. Here, we assume that the values of p and q can' be predetermined. Note that the measured speech signal, y_k, depends of the input, u_k. The excitation, u_k, is usually considered to be white Gaussian noise. We must estimate the input excitation, u_k, so that the ARMA parameters can be estimated accurately from y_k. An estimation method for u_k based on the variable forgetting factor of the WRLS algorithm will be given later. For the present we assume that an estimate for u_k is available.

Let us define a parameter vector, θ_k, its estimate, $\hat{\theta}_k$,and a data vector, Φ_k, by the following equations :

$$\theta_k{}^t = [a_1(k), \ldots, a_p(k), b_1(k), \ldots, b_q(k)] \qquad (2)$$

$$\hat{\theta}_k{}^t = [\hat{a}_1(k), \ldots, \hat{a}_p(k), \hat{b}_1(k), \ldots, \hat{b}_q(k)]$$

$$\Phi_k{}^t = [-y_{k-1}, \ldots, -y_{k-p}, u_{k-1}, \ldots, u_{k-q}]$$

where the superscript t denotes transpose, and $\hat{a}_i$ and $\hat{b}_i$ are the estimated ARMA parameters, respectively. Using (2) the speech signal, y_k, and its estimate, $\hat{y}_k$, may be expressed as

$$y_k = \Phi_k{}^t\theta_k + u_k \qquad (3)$$

$$\hat{y}_k = \hat{\Phi}_k{}^t\hat{\theta}_k + \hat{u}_k$$

Let r_k be the residual error of the ARMA process, namely,

$$r_k = y_k - \Phi_k{}^t\hat{\theta}_k \qquad (4)$$

Then a weighted least square criterion(cost function) can be defined as the exponentially weighted sum of squares of the residual errors;

$$V_k(\theta) = \sum_{i=1}^{k} \lambda^{k-i} r_i{}^2 \ \text{ or } \ V_k(\theta) = \lambda^{k-i}(y_i - \phi_i{}^t\hat{\theta}_i)^2 \qquad (5)$$

where λ is weighting(forgetting) factor, which at this point is assumed to be constant during the process of adaptation. This gives more weight to the most recent errors.

The least squares solution can be obtained by minimizing the cost function with respect to the parameter vector θ_k. A recursive least square algorithm for estimating the parameter vector θ_k can be found elsewhere [8][9]. We summarize these equations below:

Prediction error : $\xi_k = y_k - \Phi_k{}^t\hat{\theta}_{k-1}$ (6)

Gain update : $K_k = P_{k-1}\Phi_k[\lambda 1 + \Phi_k{}^t P_{k-1}\Phi_k]^{-1}$

Parameter update : $\hat{\theta}_k = \hat{\theta}_{k-1} + K_k\xi_k$

Covariance matrix : $P_k = \lambda^{-1}[P_{k-1} - K_k\Phi_k{}^t P_{k-1}]$

Input estimate : $\hat{u}_k = y_k - \phi_k{}^t\hat{\theta}_k$

B. Adaptive WRLS-VFF-VT Algorithm

For a locally stationary speech production process, the residual error r_k in (4) will indicate the state of the estimator at each instant k. If the error is small, then the forgetting factor λ should be near unity, allowing the adaptive algorithm to use most of the previous information in the signal. This yields accurate estimates of the various parameters. If, on the other hand, the error is large, then a small λ will decrease the weighting of the error, thereby shortening the effective memory length of the estimation process. This allows the parameters to be adjusted with the most recent data, and will reduce the error. A procedure to achieve the proper weighting by choosing the appropriate λ is introduced here. The weighted sum of the squares of the residual error can be expressed recursively as[10]

$$V_k(\theta) = \lambda V_{k-1}(\theta) + \xi_k^2(1-\phi_k^t K_k) \tag{7}$$

A strategy for calculating λ_k may be defined by requiring $V_k(\theta)$ to be constant such that

$$V_k(\theta) = V_{k-1}(\theta) = V_1(\theta) \tag{8}$$

In other words, the forgetting factor will compensate at each step k for the new error information in the latest measurement, thereby insuring that the estimation is always based on the same error information. Thus from (7) by setting $\lambda = \lambda_k$, we have

$$\lambda_k = 1 - \frac{\xi_k^{\ 2}}{V_1(\theta)}[1 - \Phi_k{}^t K_k] \tag{9}$$

Consequently, the WRLS-VFF algorithm can be specified by a set of equations similar to those for the WRLS algorithm in (6), but with the constant weighting factor, λ, replaced by λ_k as shown in (9).

C. WRLS-VFF-VT Algorithm with Input Estimation

The input excitation u_k to a speech production process can be either pulse trains for voiced sounds or white noise for fricatives. Therefore a general expression of the input sequence is given as follows:

$$u_k = u_k{}^p + u_k{}^w \tag{10}$$

where $u_k{}^p$ represents the pulse input and $u_k{}^w$ is the white noise input.

We introduce a weighted least squares criterion(cost function) to estimate the ARMA parameters based on the estimation error e_k(different from the residual error r_k), namely from (3) we have

$$e_k = y_k - \hat{y}_k = y_k - \phi_k{}^t \hat{\theta}_k - \widehat{u_k} \tag{11}$$

Thus the cost function $V_k(\theta)$ is expressed as

$$V_k(\theta) = \sum_{i=1}^{k} \lambda^{k-1} (y_1 - \phi_i{}^t \widehat{\theta_i} - \widehat{u_i})^2 \tag{12}$$

From (10) we have $\sum \widehat{u_k} = \sum \widehat{u_k}^w + \sum \widehat{u_k}^p$, and under the theorem of ergodicity, for large k, these sums are expectations[10], which indicates that when the input to the reference model is white, the prediction error produced by an optimal prediction (e.g., WRLS algorithm) is also white. Thus, for a zero mean white noise input, we have $E[\widehat{u_k}^w] = E[u_k{}^w] = 0$. Then (12) can be rewritten as

$$V_k(\theta) = \sum_{i=1}^{k} \lambda^{k-1} (y_1 - \phi_i{}^t \widehat{\theta_i} - \widehat{u_i}^p)^2 \tag{13}$$

The parameter vector θ_k that minimizes the cost function $V_k(\theta)$ can be obtained by setting the parallel derivative of V with respect to θ equal to zero, which gives the recursive equation to update $\widehat{\theta_k}$ as

$$\widehat{\theta_k} = \hat{\theta}_{k-1} - K_k (y_k - \phi_k{}^t \hat{\theta}_{k-1} - \widehat{u_k}^p) \tag{14}$$

Several methods have been proposed to estimate the input u_k in the recursive ARMA parameter estimation algorithm. Morikawa and Fujisaki and Friedlander used the estimated residual error r_k at the instant k as the estimated input u_k, namely [11][12]

$$\widehat{u_k} = r_k = y_k - \phi_k{}^t \widehat{\theta_k} \tag{15}$$

This method is based on the fact that the driving source to the ARMA process is a pure white noise. Miyanaga et al. used the ratio of the variance of zero-mean white noise at k and k+1 to estimate the forgetting factor(FF)[8]. For application to speech signal analysis when the input contains both white noise and pulses, we need to estimate both of these signals to obtain more accurate parameter estimates.

The proposed input estimation method uses the FF as a reference to examine the input condition. This can reduce the system complexity since only one adaptive algorithm is used instead of two as in [8]. Moreover, the FF can be obtained from the adaptive process. Thus no extra calculations are required.

From (9), it has been shown that the increase in the prediction error results in a decrease in λ_k. A small value of λ_k indicates that the input has an abrupt change (e.g. pulse). Hence, we can determine the time of occurrence of a pulse by determining the instant at which λ_k falls below a threshold value λ_{thr}. A strategy for choosing the threshold value λ_{thr} may now be defined by letting

$$E_k = 1/M \sum_{i=0}^{M-1} \lambda_{k-i} \qquad (16)$$

If $E_k < 0.9$, *then* $\lambda_{thr} = 0.99 * E_k$

If $E_k > 0.9$, *then* $\lambda_{thr} = 0.9 * E_k$

If $\lambda_{thr} > \lambda_{min}$, *then* $\lambda_{thr} = \lambda_{min}$

The magnitude of the pulse can be approximately given by the prediction error ξ_k at the estimated time of the input pulse [8]. For white noise input, the λ_k is close to unity upon convergence [10]. Under this condition the residual error r_k of the adaptive process can be used as the estimate of the white noise input u_k^w as indicated in Morikawa's method [11]. Consequently, the implementation of the WRLS-VFF-VT algorithm with input estimation is given in Table I. Figure 1 shows the overall block diagram to estimate ARMA parameters.

Prediction error : $\xi_k = y_k - \widehat{\Phi}_k^{\,t}\, \hat{\theta}_{k-1}$

Gain : $K_k = P_{k-1}\widehat{\Phi}_k[\lambda_{k-1} + \widehat{\Phi}_k^{\,t} P_{k-1}\widehat{\Phi}_k]^{-1}$

Forgetting Factor : $\lambda_k = 1 - \xi_k^2(1 - \widehat{\Phi}_k^{\,t} K_k)^2 / V_1(\theta)$

Threshold : $V_k = 1/M \sum_{i=1}^{M-1} \lambda_{k-i}$

If $E_k < 0.9$, then $\lambda_{thr} = 0.99 * E_k$

If $E_k > 0.9$, then $\lambda_{thr} = 0.9 * E_k$

If $\lambda_{thr} > \lambda_{min}$, then $\lambda_{thr} = \lambda_{min}$

Input estimate:

a) If $\lambda_k < \lambda_{thr}$, then the input is a pulse and

$$\widehat{u}_k^{\,w} = 0$$
$$\widehat{u}_k = \widehat{u}_k^{\,p} = y_k - \widehat{\Phi}_k^{\,t}\, \hat{\theta}_{k-1}$$

b) If $\lambda_k > \lambda_{thr}$, then the input is white noise and

$$\widehat{u}_k^{\,p} = 0$$
$$\widehat{u}_k = \widehat{u}_k^{\,w} = \xi_k(1 - \widehat{\Phi}_k^{\,t} K_k)$$

Parameter : $\hat{\theta}_k = \widehat{\theta_{k-1}} + K_k e_k$

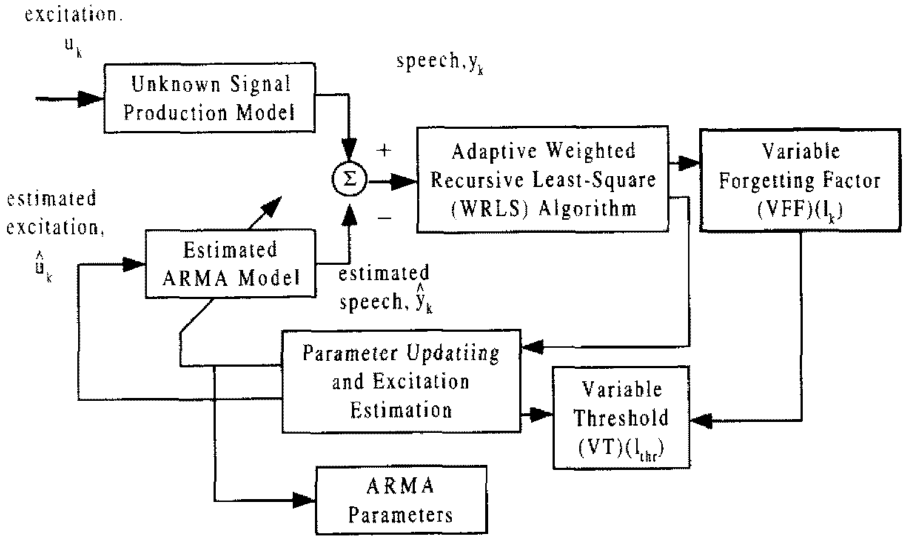
Covariance matrix : $P_k = \lambda_k^{-1}[P_{k-1} - K_k \widehat{\Phi}_k^{\,t} P_{k-1}]$

<Table 1> Adaptive WRLS-VFF-VT Algorithm with Input Estimation

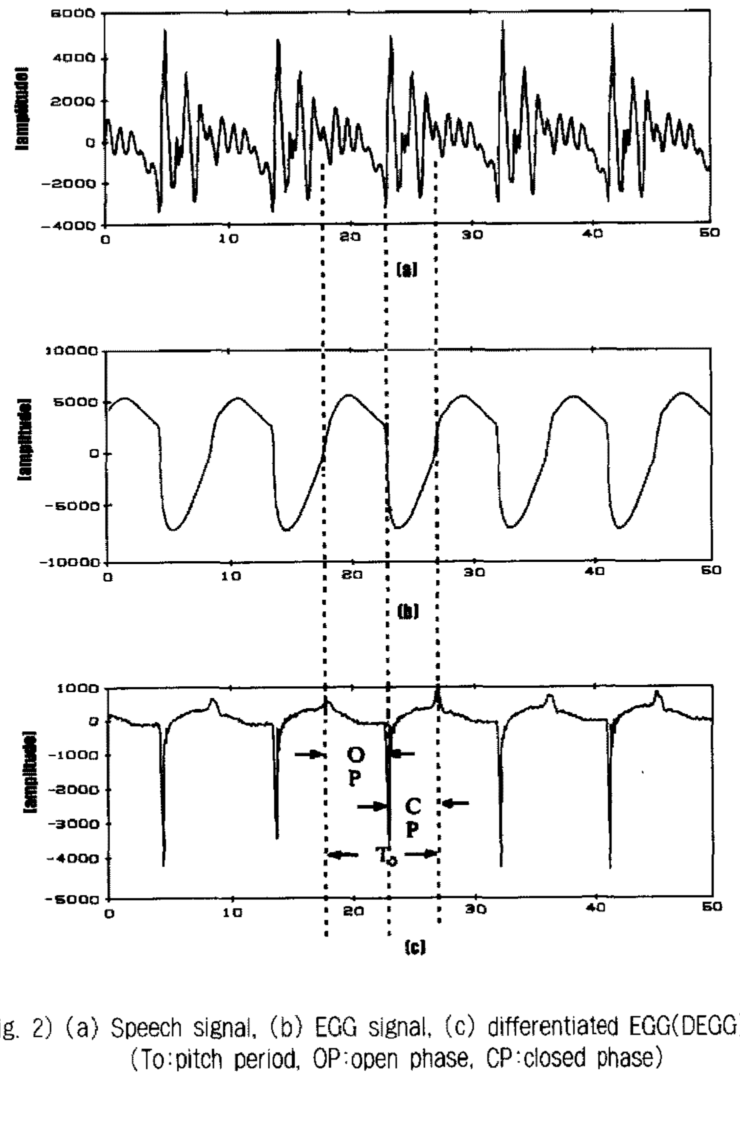
3. Proposed Pitch Detection Algorithms

A. Modified EGG Based PDA with variable threshold

The EGG(ElectroGlottoGraphic) based PDA was first proposed by Krishnamurthy and Childers in 1986 [3]. The EGG is a periodic signal with two zero crossings per period during the voiced segments. The pitch period can be estimated as the time duration between two successive "invariant" features in the EGG. A synchronized, differential EGG(DEGG) signal is used to locate the closed glottal phase and to detect the pitch period. Figure 2 shows an example of EGG and DEGG waveform with the synchronized speech signal using simple threshold as in [3].

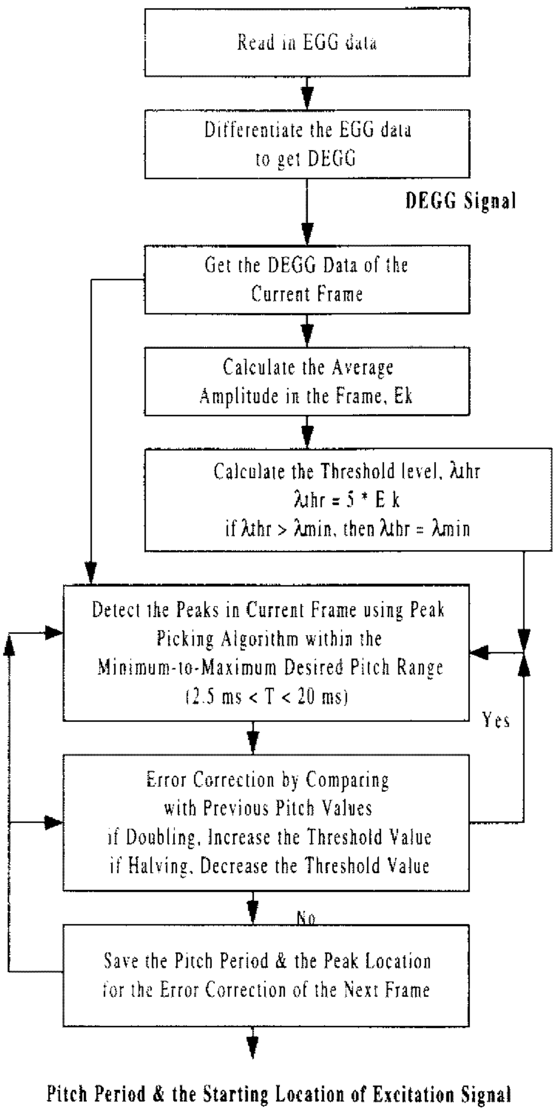


(Fig. 1) Block Diagram of the WRLS-VFF-VT algorithm for the ARMA parameter estimation



(Fig. 2) (a) Speech signal, (b) EGG signal, (c) differentiated EGG(DEGG) signal
(To:pitch period, OP:open phase, CP:closed phase)

As can be seen in this figure, it is not difficult to obtain the pitch period, the glottal opening and closing instants from the EGG signal. While the glottal opening occurs relatively slowly, glottal closure is associated with a rapid reduction in tissue impedance and thus shows a large negative excursion in the EGG signal. Therefore, a simple threshold may give the missing of the peak position in the beginning of a sentence and in the mixed sounds and in the weak sounds. Moreover, for a pathological speech mode such as vocal fry which has two or three opening/closing pulses in a glottal cycle, it remains difficult to detect the pitch period with the EGG signal. To solve this difficulty, the variable threshold is used in this paper.
Figure 3 shows the block diagram of the modified EGG based pitch detection algorithm with variable threshold. The EGG signal is divided into frames, each frame consisting of 300 points. Successive frames overlap by 200 points. The maximum EGG amplitude and the EGG zero crossing rate in the frame are used to classify the frame as voiced or unvoiced. In the voiced frames, the average amplitude in a current frame is computed for the variable threshold of the peak picking using the DEGG. The opening and closing instants in the frame are located using the DEGG. Two frames of delayed pitch information are retained for the pitch error detection and correction.



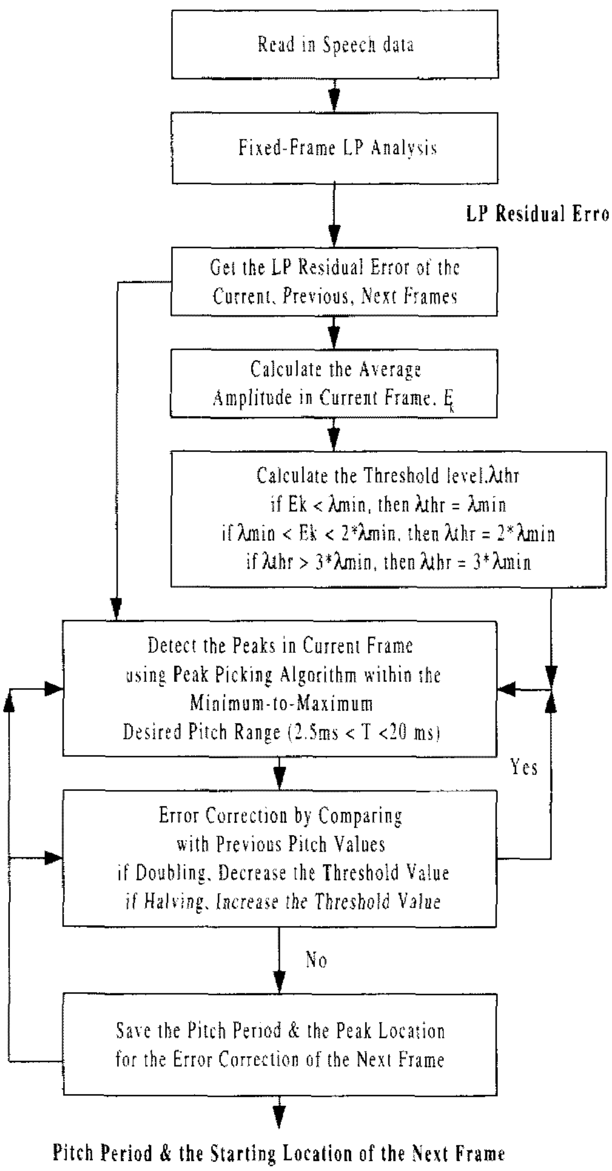
(Fig. 3) Block Diagram of EGG based pitch detection algorithm

The EGG signal is immune to surrounding acoustic disturbance, and provides a robust pitch detector. Moreover, the EGG can also be used to isolate individual periods of the speech waveform such as the interval of glottal closure.

B. Modified LP Error Based PDA with Variable Threshold

In 1973, Markel[4] proposed a LP Error based pitch estimation with fixed threshold method that employs the autocorrelation function of the LP error signal. Markel showed that the LP error signal for a voiced speech waveform is characterized by peaked pulses separated by the pitch periods. These peaked pulses represent the main excitations to the LP vocal tract filter. As can be seen in figure 2, the main excitation pulses in the LP error function consistently match the negative peaks of the DEGG signal, which were located very close to the instants of glottal closure. This method is capable of providing accurate pitch periods, but loses all information about the absolute position of the glottal excitation. Therefore, we use the variable threshold in this study.

A block diagram of the modified LP Error based method is shown in figure 4.



(Fig. 4) Block Diagram of LP Error based pitch detection algorithm

A pitch-asynchronous (fixed frame) LP analysis is performed on the input speech signal. For a voiced speech signal, the LP error function is characterized by a pulse train with the appropriate pitch period. The locations of these pulses are detected by a peak-picking method with variable threshold and are used as indicators of glottal closure.

The LP residual error signal is divided into frames, each frame consisting of 100 points. The analysis frame is overlapped by 100 points of previous and next frames, thereby totally 300 points. In the voiced frames, the average amplitude in a current frame is computed for the threshold of peak picking. Two frames of delayed pitch information are retained for the pitch error detection and correction. The LP error based method with variable threshold for pitch detection is capable of providing the pitch on a period by period basis with no extra auxiliary signal(EGG).

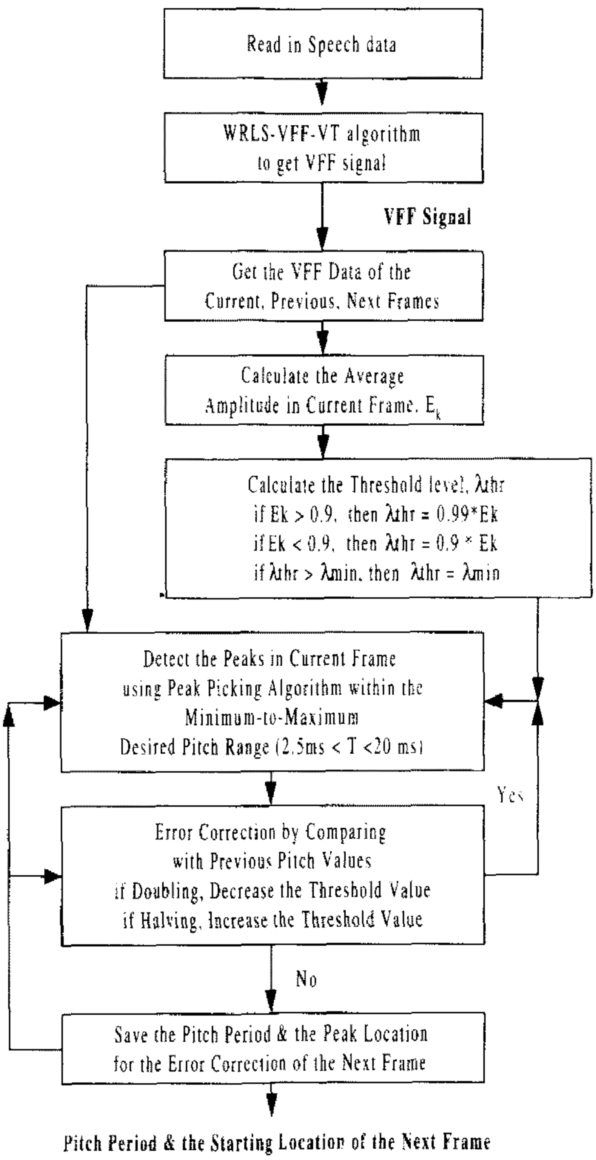
C. VFF Based PDA

The VFF based PDA from WRLS-VFF-VT algorithm described in section II uses the VFF to identify the glottal closure points, which correspond to the instants of occurrence of the main excitation pulses for voiced speech. The smallest value of λ_k in the WRLS-VFF-VT algorithm consistently matches the negative peaks of the DEGG signal, thereby indicating the instant of glottal closure in figure 2.

A block diagram of the VFF based method for the pitch information is shown in figure 5.

During the WRLS-VFF-VT analysis, the λ_k can be obtained sequentially(sample-by-sample). The location of the excitation pulses and the estimation of the pitch period are detected by using a threshold value of λ_{min} and comparing it with each λ_k.

A variable threshold is used in the same way as in the LP error based method. If a peak crosses the variable threshold, its location becomes the pitch period candidate. Otherwise the frame is defined as unvoiced (i.e., pitch period = 0). An attempt at error correction is made by changing the threshold of the pitch period. The pitch period estimation occurs for the range 2.5 msec to 15.5 msec. The VFF based method for pitch detection is capable of providing the pitch on a period by period basis with no extra auxiliary signal (EGG).



(Fig. 5) Block Diagram of VFF Based pitch detection algorithm

4. Performance Evaluation

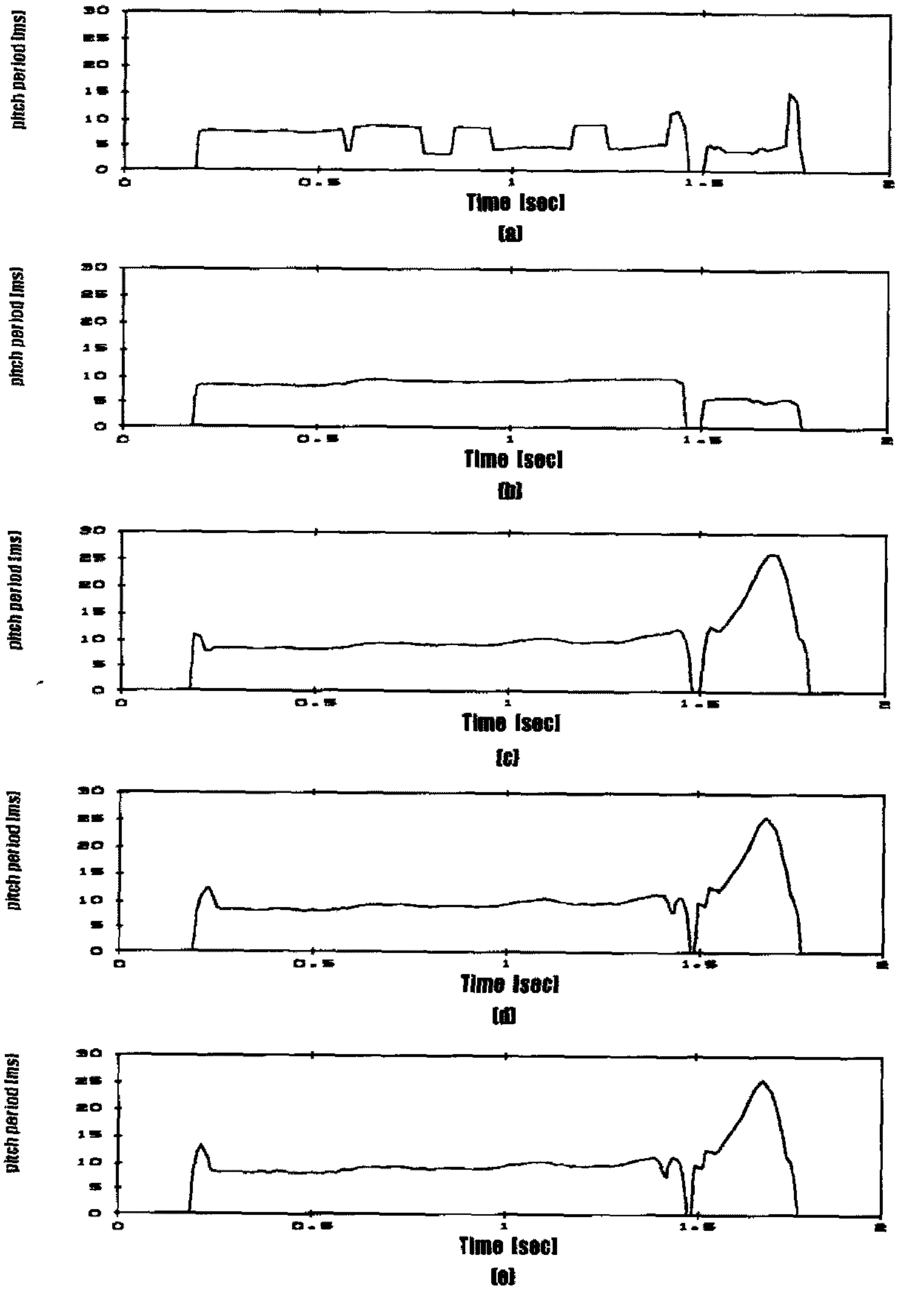
In order to evaluate the performance(robustness) of our pitch detector, we compare the results for pitch estimation for the speech based SIFT algorithm with proposed methods in natural human speech as well as in pathologic human speech. We may use the EGG based method as a reference.

Figure 6 shows the pitch contours for the utterance "We were away a year ago" spoken by a male speaker.

Figure 6 (c),(d) show the results of the EGG based, LP error based, and (e) shows the results for the proposed VFF based algorithms, respectively. The pitch contours estimated from the proposed three methods are virtually identical and show the fine detail of the intonation changes. However, the result for the SIFT algorithm in Figure 6 (a) has the error caused by the pitch period doubling or tripling or halving, and also missed some fine detail of the intonation changes. Using the modified SIFT algorithm of variable analysis frame size, the errors of the pitch period doubling and halving were corrected as shown in Figure 6 (b). But the modified SIFT algorithm did not perform well for the aperiodic voice vibration contained in the end of the sentence.

For the pathological speech mode, pitch contours for sustained vowel "i" spoken by a pathologic speaker, vocal fry, are shown in figure 7.

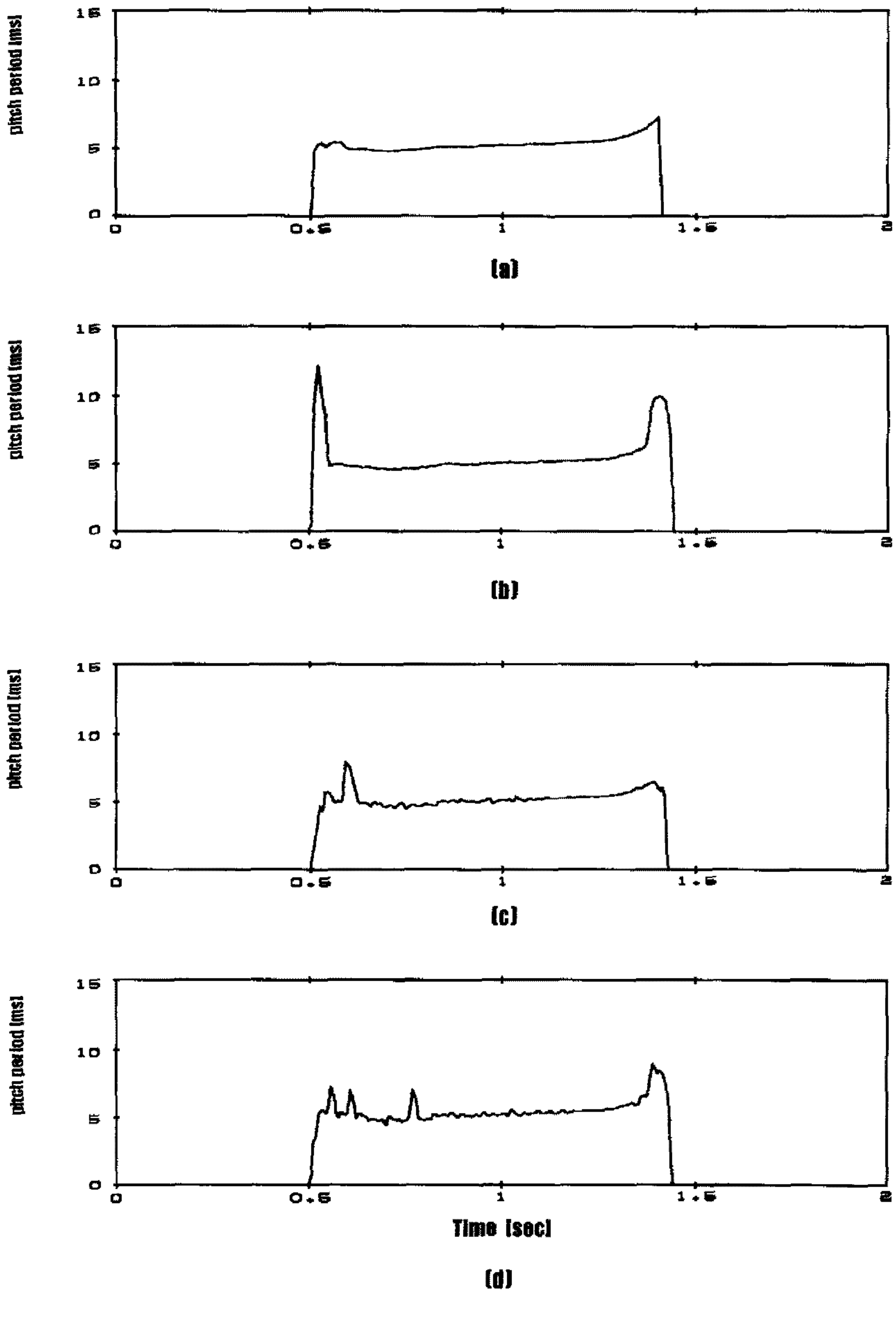
In figure 7 (a), the contour derived from the modified SIFT algorithm is smooth, but some fine detail is missed. Unlike the proposed methods, the modified SIFF algorithm uses an average pitch estimate over several periods, so the result provides the similarity over the adjacent pitch periods. The result from EGG based algorithm in figure 7 (b) shows the details of the intonation and two abrupt changes in the start and in the end of the signal. This is caused by the incomplete glottal opening/closing in the beginning and ending area. The LP errors based and the VFF based algorithms produce contours that have ripples, spurious values in the middle of contours, and the fine detail of the intonation changes as shown in Figure 7 (c) and (d).



(Fig. 6) Pitch contours for a sentence "we were away a year ago" using:
(a) SIFT, (b) modified SIFT, (c)modified EGG based,
(d) modified LP error based, and (e) VFF based methods

5. Summary and Conclusion

New time domain pitch detection algorithms, modified EGG-based, modified LP Error- based, and WRLS-VFF-VT based methods that provide pitch estimates on a period by period basis, have been described. In comparison with the SIFT algorithm, it is shown that new algorithms have the performance which are closest to that obtained using the EGG for the speech of normal and pathologic speakers. If



(Fig. 7) Pitch contours from four algorithms for sustained vowel "i" spoken by a pathologic speaker, vocal fry: (a) modified SIFT, (b) modified EGG based, (c) modified LP error based, (d) VFF based

the EGG based method is considered as a reference, the proposed pitch detectors are very reliable in quasi-periodic as well as in aperiodic speech signals. The "pitch smearing" effect inherent in speech signal based methods is avoided, and pitch values are available on a period by period basis. This can have important applications for several problems where accurate pitch contour estimates are desired [13],

References

[1] Markel, J.D., "The SIFT algorithm for fundamental frequency estimation", IEEE Trans on ASSP, Vol.AU-20, pp.367-377, 1972.

[2] Childers, D. G., Skinner, D.P. and Kemerait R.C, "The cepstrum: A guide to processing", Proc. of IEEE, Vol.65, No.10, pp.1428-1443, 1977.

[3] Krishnamurthy A.K., and Childers D.G., "Two-channel speech analysis", IEEE Trans. on ASSP, Vol.34, No.4, pp.730-743, 1986.

[4] Atal B.S. and Hanauer S.L., "Speech analysis and synthesis by linear prediction of the speech wave", Journal of Speech and Hearing Research, Vol.50, pp.637-655, 1975.

[5] Akaike H., "A new look at the statistical model identification", IEEE Trans on Automatic Control, Vol.AC-19, pp.716-723, 1974.

[6] Kay, "Modern Spectrum Estimation", Prentice-Hall Inc., Englewood Cliffs, NJ 1987.

[7] Marple S.L., "Digital Spectral Analysis with Applications", Prentice-Hall Inc., Englewoord Cliffs, NJ 1987.

[8] Miyanaga Y. Miki, N. Nagai, and Hitori K., "A speech analysis algorithm which eliminates the influence of pitch using the model reference adaptive system", IEEE Trans on ASSP, Vol.30, No. 1, pp.88-95, 1982

[9] Soderstrom T., and SToica P. "System Identification", Prentice-Hall Inc., Englewood Cliffs, NJ 1989.

[10] T.R. Fortescue, L.S. Kershenbaum and B.E. Ydstic, "Implementation of self-regulators with variable forgetting factors", Automatica, Vol.17, pp.831-835, 1981.

[11] Morikawa H. and Fujisake H., "Adaptive analysis of speech based on a pole-zero representation", IEEE Trans. on ASSP, Vol.30, No.1, pp.77-87, 1982.

[12] Friedlander B., "A recursive maximum likelihood algorithm for ARMA spectral estimation", IEEE Trans. on Information Theory, Vol.IT-28, No.4, pp.639-646, 1982

[13] Hess W.J., "Algorithms and devices for pitch determination of speech signals", Phonetica, Vol. 39, pp.219-240, 1982.

이 교 식

1982년 경북 대학교 전자공학과 졸업(학사)

1984년 경북 대학교 전자공학과 졸업(석사)

1984년~1989년 국방과학연구소 연구원

1992년 (미) University of Florida, 전자공학과 공학 박사

1993년~1994년 삼성전자 이동통신기술센터, 수석연구원

1994년~1995년 한국해양대학교 전파공학과, 전임강사

1995년~1997년 (주)서한전자, 대표이사겸 멀티미디어 부설연구소 소장

1997년~1998년 (주)삼테크, 뉴미디어본부장 겸 부설연구소 소장

1998년~현재 한세대학교 정보통신학과 조교수

관심분야 : Digital Speech Processing, Digital Communication

박 규 식

1986년 (미) Polytechnic University, 전자공학과 학사

1988년 (미) Polytechnic University, 전자공학과 석사

1993년 (미) Polytechnic University, 전자공학과 공학박사

1994년~1995년 삼성전자 마이크로사업부, 선임연구원

1995년~1996년 한국 해양대학교 전파공학과, 전임강사

1996년~현재 상명대학교 컴퓨터·정보통신학부 조교수

관심분야 : Digital signal and image processing, Digital communication systems