

성대의 비선형 2-mass 디지털 모델 구현

Implementation of nonlinear two-mass vocal folds digital model

이희승*, 정명진**

Hui Sung Lee, Myung Jin Chung

Abstract - The vocal folds play an important role to produce glottal pulse which is an essential factor of phonation. There have been some models which implement the vocal folds' dynamics, such as one-mass model, two-mass model, multi-mass model and ribbon model. Among them, this paper uses nonlinear two-mass model, which is simple structure and produces similarly real glottal pulses and vocal folds' vibration, to realize vocal folds digital model. The pattern of movements in vocal folds will be shown in this paper by using vocal folds digital model. It will be verified how initial position of vocal folds, variation of tension and change of lung pressure influences vibration and glottal pulses.

Key Words : vocal folds, digital model, two-mass, articulatory synthesis

1. 서론

사람이 말소리를 생성하기 위해서는 발동(Initiation), 발성(phonation), 조음(articulation)과 같은 세 가지 과정을 거쳐야 한다[1]. 이 중 발성은 발동에 의해 공급된 기류를 적절히 조절하는 과정으로, 대부분 성대 진동에서 기인한다. 이와 같이 발성의 중요한 역할을 하는 성대에 대해 모델링하고 성대로부터 나오는 성문 파동(glottal pulse)의 구현은 조음 음성 합성법에 있어서 핵심 기술 중에 하나이다[2]. 성대의 모양 및 길이, 두께의 차이 등이 음색을 결정짓는 요소가 되고, 베르누이(Bernoulli) 효과에 의해 성대가 떨리게 된다[1][3][4]. 성문 파동을 특정한 함수로 표현하여 조음음성 합성을 시도하기도 하지만[2], 성대의 떨림을 관찰하고 모델링하여 성문 파동을 구현하기 위해 one-mass 모델에서 출발하여 two-mass 모델, multi-mass 모델, ribbon 모델에 이르기까지 다양한 형태의 모델이 제시되어 왔다[5]-[8]. 그러나 모델의 자유도가 높아질수록 실제 성대 움직임에 더 근접한 효과를 얻을 수 있지만, 상대적으로 계산량이 많아져서 실제 구현시 어려운 점이 있다[7].

본 논문에서는 간단한 구조를 가지면서도 실제 성대 떨림 및 성문 파동 모양에 근접한 결과를 보이는 two-mass 모델 [6]-[8]을 연구하고 디지털 시스템으로 구현해서 성대의 떨림의 유형을 미리 관찰함을 목표로 하였다. 구현된 디지털 모델을 이용해서 다양한 초기값의 변화, 긴장도의 변화, 폐 압력 변화 등을 모의 실험하여 각각의 영향을 관찰하고, 각 조건에 따른 성대 떨림의 특징을 분석하였다.

2. 성대의 비선형 two-mass 모델

Ishizaka가 제안한 방식을 바탕으로 성대를 모델링하여 그림으로 나타내면 그림 1과 같다. m_1 과 m_2 은 성대의 질량을

표현한 것이고, s_1 과 s_2 는 등가 스프링을, r_1 과 r_2 는 등가 점성 저항을 나타낸다. 이와 같은 모델은 다음과 같은 몇 가지 가정으로부터 시작한다[6][7].

- (1) 상위의 성대 질량(m_2)은 하위의 질량(m_1)보다 작다.
- (2) 성대의 두 개 질량(m_1, m_2)은 한 쪽 방향으로의 운동(수평운동, lateral motion)만 한다.
- (3) 성대가 이루는 면적의 모양은 직사각형이다.
- (4) 성문에서의 공기의 흐름은 한 쪽 방향(성대의 수직 방향)으로 흐른다. (quasi-steady)
- (5) m_1 과 m_2 사이에도 linear stiffness가 존재한다.

성대 모델을 구현하기 위해 사용되는 특정 상수는 다음과 같다.

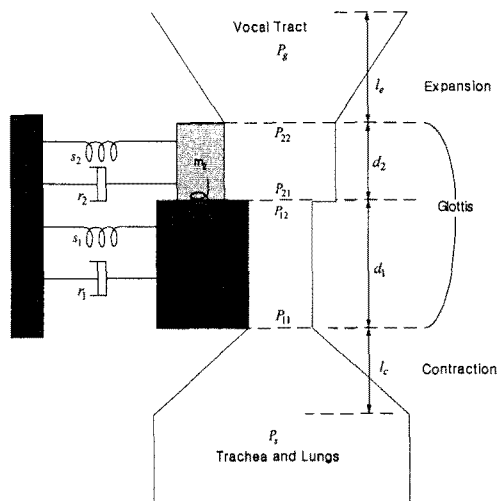


그림 1. 성대의 two-mass 모델을 측면(단면)에서 본 것
 μ : viscosity. 상온(37°C)에서는 $1.86 \times 10^{-4} \text{ dyne} \cdot \text{s/cm}^2$
 ρ : air density, 상온에서는 $1.14 \times 10^{-3} \text{ g/cm}^3$

저자 소개

* 正 會 員 : KAIST 전기및전자공학전공 박사과정

** 正 會 員 : KAIST 전기및전자공학전공 교수

공기가 흐르면서 생성되는 압력을 구하기 위해서는 성대의 단면적을 먼저 구해야하고, 단면적에 따른 공기 부피 속도 U_g 를 구해야 한다. 성대의 각 부분의 압력은 구조에 따라 다른 형태의 압력 요소가 존재하고, m_1 과 m_2 에 영향을 주는 압력을 개별적으로 구해야 한다. 주어진 압력과 스프링의 운동 특성을 이용해서 성대의 움직임을 표현하는 미분방정식을 구할 수 있다. 스프링의 운동 특성과 m_1 과 m_2 에 영향을 주는 압력은 구조적으로도 비선형적인 요소를 가지고 있으며, 성문이 열리고 닫히는 상황에 따라서도 각기 다른 특성을 나타낸다. two-mass 모델의 미분 방정식 유도과정과 특성은 참고 문헌 [5][6][7]을 참고하고, 이를 요약 정리하면 다음의 식과 같다.

$$u_1 = U_g/A_{g1}, \quad u_2 = U_g/A_{g2}, \quad P_s - P_{11} = 1.37 \frac{\rho u_1^2}{2}$$

$$P_{11} - P_{12} = 12\mu u_1 d_1 \left(\frac{l_g}{A_{g1}} \right)^2, \quad P_{12} - P_{21} = \frac{\rho}{2} (u_2^2 - u_1^2) \quad (1)$$

$$P_{21} - P_{12} = 12\mu u_2 d_2 \left(\frac{l_g}{A_{g2}} \right)^2, \quad P_{22} - P_1 = -\rho u_2^2 \frac{A_{g2}}{A_1} \left(1 - \frac{A_{g2}}{A_1} \right)$$

m_1 과 m_2 에 가해지는 각각의 압력 P_{m1} 과 P_{m2} 를 다음과 같이 구할 수 있다.

$$P_{m1} = \frac{P_{11} + P_{12}}{2} = P_s - 1.37 \frac{\rho u_1^2}{2} - 6\mu u_1 d_1 \left(\frac{l_g}{A_{g1}} \right)^2 \quad (2)$$

$$P_{m2} = \frac{P_{21} + P_{22}}{2}$$

$$= P_{m1} - 6\mu u_1 d_1 \left(\frac{l_g}{A_{g1}} \right)^2 - 6\mu u_2 d_2 \left(\frac{l_g}{A_{g2}} \right)^2 - \frac{\rho}{2} (u_2^2 - u_1^2)$$

m_2 에 전달되는 힘과 거리(x_2)의 관계식은 다음과 같다.

$$m_1 \frac{d^2 x_1}{dt^2} + r_1 \frac{dx_1}{dt} + f_1(x_1) + k_c(x_1 - x_2) = F_1$$

$$m_2 \frac{d^2 x_2}{dt^2} + r_2 \frac{dx_2}{dt} + f_2(x_2) + k_c(x_2 - x_1) = F_2 \quad (3)$$

$f_j(x_j)$ 는 다음과 같이 x_j 에 따라 형태로 표현된다.

$$\text{if } x_j > -x_{j0}, \text{ then } f_j(x_j) = k_j(x_j + \eta_{kj}x_j^3)$$

$$\text{if } x_j \leq -x_{j0}, \text{ then} \quad (4)$$

$$f_j(x_j) = k_j(x_j + \eta_{kj}x_j^3) + h_j \{ (x_j + x_{j0}) + \eta_{hj}(x_j + x_{j0})^3 \}$$

식 (4)와 함께 F_j 도 x_1 과 x_2 의 위치에 따라 4가지 경우의 수가 있어서[6], 구현된 모델의 비선형성 특징을 제공하고 있다.

3. 디지털 시스템으로 구현

주어진 미분방정식 (3)를 디지털 시스템으로 구현하기 위해 backward approximation을 사용해서 정리한다. 단, 상태 방정식으로 표현하기 위해 $z_1 = x_1$, $z_3 = x_2$ 로 두고 정리하면 다음과 같다.

$$z_1[n] = z_1[n-1] + Tz_2[n-1]$$

$$z_2[n] = \frac{T}{m_1} \{ -f_1(z_1[n-1]) - r_1 z_2[n-1] - k_c(z_1[n-1] - z_3[n-1]) + F_1[n-1] \} + z_2[n-1]$$

$$z_3[n] = z_3[n-1] + Tz_4[n-1] \quad (5)$$

$$z_4[n] = \frac{T}{m_2} \{ -f_2(z_3[n-1]) - r_2 z_4[n-1] - k_c(z_3[n-1] - z_1[n-1]) + F_2[n-1] \} + z_4[n-1]$$

식 (2)와 식 (3), (4)의 경우는 미분방정식이 아니므로 근사

화 없이 그대로 사용하고, 성문 파동의 결과값인 U_g 는 다음과 같은 식을 이용해서 구할 수 있다.

$$U_g[n] = \frac{-b + \sqrt{b^2 - ac}}{a} \quad (6)$$

$$a = \rho \left\{ \frac{0.685}{A_{g1}^2[n]} + \frac{1}{2} \left(\frac{1}{A_{g2}^2[n]} - \frac{1}{A_{g1}^2[n]} \right) \right\} - \frac{1}{A_{g2}[n]A_1} \left(1 - \frac{A_{g2}[n]}{A_1} \right)$$

$$b = 6\mu l_g^2 \left(\frac{d_1}{A_{g1}^3[n]} + \frac{d_2}{A_{g2}^3[n]} \right)$$

$$c = -P_s + P_1$$

4. 실험 결과

실험 전에 상수로 고정시켜야 하는 값과 변수를 구분해야 하는데, 본 논문에서는 성대의 긴장도(tension) Q 와 초기 위치(x_{10} , x_{20}), 폐의 압력 P_s 의 변화에 대한 성대 떨림의 유형을 관찰한다. 실험에 사용된 기본적인 상수 값은 다음과 같다[4]-[6][8].

$$P_s = 8 \text{ cmH}_2\text{O} = 8 \times 980 \text{ dynes/cm}^2, \quad l_g = 1.4 \text{ cm},$$

$$d_1 = 0.25/Q \text{ cm}, \quad d_2 = 0.05/Q \text{ cm},$$

$$\eta_{k1} = \eta_{k2} = 100, \quad \eta_{h1} = \eta_{h2} = 500,$$

$$k_1 = 80000 Q \text{ dyn/cm}, \quad k_2 = 8000 Q \text{ dyn/cm},$$

$$k_c = 25000 Q \text{ dyn/cm},$$

$$m_1 = 0.125/Q \text{ g}, \quad m_2 = 0.025/Q \text{ g}, \quad h_1 = 3k_1, \quad h_2 = 3k_2$$

4.1 초기 위치(x_{j0})에 따른 변화

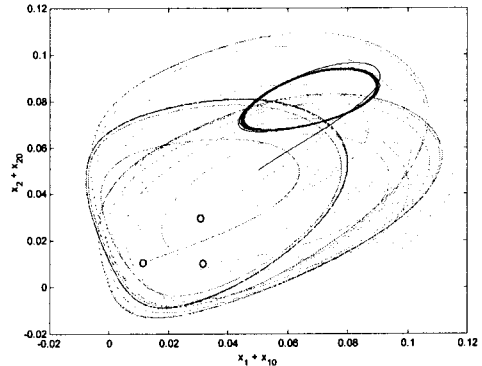


그림 2. 초기 위치 x_{j0} 에 따른 x_j 의 phase portraits. $P_s = 8 \text{ cmH}_2\text{O}$, $Q = 0.8$

성대의 초기 위치에 따라서 성대가 진동할 수도 있고 안 할 수도 있다. 직관적으로 성대가 너무 붙어 있거나, 너무 떨어져 있으면 진동하지 않을 것을 예상할 수 있는데[4], 실험 결과에서도 이를 확인할 수 있다. 실험 시 $P_s = 8 \text{ cmH}_2\text{O}$ $Q = 0.8$ 를 사용하였고, 그 결과는 그림 2와 같이 phase portraits로 표현하였다. 초기값은 $(x_{10}, x_{20}) = (0.01, 0.01) \sim (0.05, 0.05)\text{cm}$ 범위에서 0.02 단위로 시행되었고, 결과적으로 $(0.01, 0.01)$, $(0.03, 0.01)$, $(0.03, 0.03)$ 에서만 닫힌 궤적이 존재하고 다른 경우에는 존재하지 않았다. 이를 통해 성대가 진동할 수 있는 적절한 성대 초기위치가 존재함을 확인할 수 있다.

4.2 성대 긴장도(Q)의 변화

성대의 긴장도는 성문 파동의 주기를 결정하므로 음높이와

관련 있다. 긴장도가 클수록 성대의 진동이 일어나기 힘들고 성문 파동의 주기가 짧아질 것을 예측할 수 있다[4][6]. 이러한 사실도 실험을 통해 확인할 수 있다. 4.1절의 실험처럼 동일한 초기위치를 설정하고 Q 를 0.4~1.0 까지 변화시키면서 성대 떨림의 유형을 관찰하였다. 실험결과 Q 가 작아질수록 성대가 진동할 수 있는 초기위치의 수가 증가하였고, 그림 3을 통해 Q 가 증가할수록 U_0 의 주기도 짧아져서 음높이 증가함을 알 수 있다.

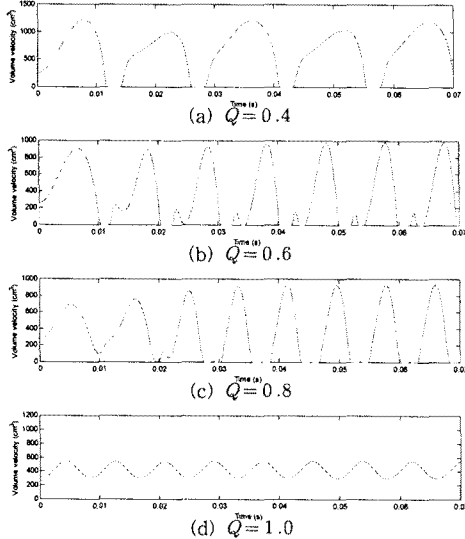


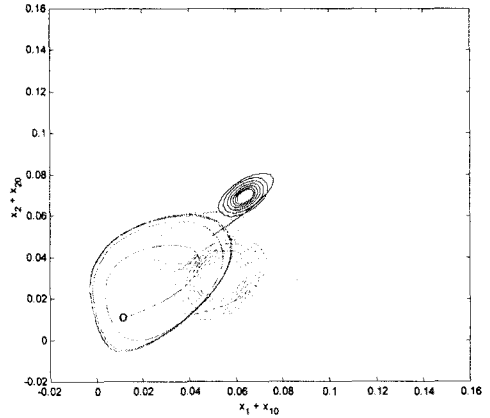
그림 3. 성대 긴장도 Q 의 변화에 따른 성대 진동의 변화. $(x_{10}, x_{20})=(0.03, 0.03)$, $P_s=8$ cmH₂O

4.3 폐 압력(P_s)의 변화

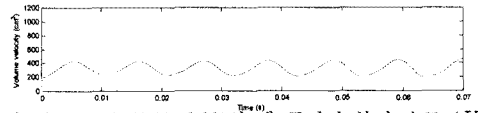
폐의 압력과 성대 진동과의 관계를 예측해 보면, 압력이 높아질수록 공기의 최대 부피 속도가 증가해서 음량과 음높이가 증가하고, 성대 진동이 가능한 초기 위치 영역도 증가할 것으로 예상할 수 있다[4]. 실험 결과에서도 압력이 증가할수록 phase portraits의 닫힌 궤적을 생성하는 초기위치 영역이 증가하고, 궤적의 크기 또한 크면서 짧은 주기를 보였다. 그림 4는 그림 2와 그림 3.(c)과 비교할 수 있도록 $P_s=5$ cmH₂O 일 때의 phase portraits와 U_0 를 나타낸 것이다. 그림 2나 그림 3.(c)의 경우에 비해 닫힌 궤적 발생 초기 위치 영역과 궤적의 크기가 감소하였고, 진동 주기는 증가하여 낮은 음높이를 나타냄을 알 수 있다.

5. 결론

본 논문에서는 기존에 제안된 여러 성대 모델 중에서 자유도가 높지 않으면서도 실제 성대의 움직임을 잘 모사하고 실제 성문 파동과 유사한 형태를 생성하는 two-mass 모델을 이용해서 디지털 시스템으로 구현하였다. 구현된 디지털 시스템으로 성대의 초기 위치, 폐의 압력 그리고 성대의 긴장도가 성문 파동에 어떠한 영향을 주는지 확인하였다. 연구 결과 성대가 진동하기 위한 적절한 초기 위치가 있는데, 성대



(a) phase portraits



(b) $(x_{10}, x_{20})=(0.03, 0.03)$ 일 때 공기의 부피 속도 (U_0)
그림 4. $P_s=5$ cmH₂O에서의 성대 위치 phase portraits와 출력 공기의 부피 속도. $Q=0.8$

진동이 일어나지 않는다. 성대의 긴장도는 음높이에 직접적인 영향을 주는데, 긴장도가 증가할수록 음높이도 증가한다. 그러나 너무 높은 긴장도는 오히려 성대 진동을 방해함을 확인하였다. 적절한 폐 압력도 성대 진동에 필요한 요소이다. 폐의 압력이 너무 낮으면 진동이 일어나는 영역이 감소하고 압력이 높아질수록 음량과 음높이가 증가함을 확인하였다. 본 연구를 기반으로 추후 조음음성합성기의 성문 파동부에 이러한 디지털 모델이 사용되어 더 자연스러운 음성합성기 구현에 도움이 될 것이다[2][3][5].

참고 문헌

- [1] 신지영, *말소리의 이해*, 한국문화사, 2000
- [2] 이희승, *조음기관 모델링을 통한 한국어 모음 음성 합성기 구현*, KAIST 석사학위논문, 2002
- [3] L. R. Rabiner and R. W. Schfer, *Digital Processing of Speech Signals*, Prentice-Hall, 1978
- [4] K. N. Stevens, *Acoustic Phonetics*, MIT Press, 1999
- [5] F. Avanzini, P. Alku, and M. Karjalainen, "One-delayed-mass Model for Efficient Synthesis of Glottal Flow", In Proc. Eurospeech Conf, 51~54, 2001
- [6] K. Ishizaka and J. L. Flanagan., "Synthesis of Voiced Sounds From a Two-mass Model of the Vocal Cords", Bell System Tech. J. vol 51, 1233~1268, 1972
- [7] I. R. Titze and B. H. Stroy, "Rules for controlling low-dimensional vocal fold models with muscle activation", J. Acoust. Soc. Am. vol 112(3), 1064~1076, 2002
- [8] J. J. Jiang and Y. Zhang, "Chaotic vibration induced by turbulent noise in a two-mass model of vocal folds", J. Acoust. Soc. Am. 112(5), 2127~2133, 2002