

ในปัจจุบันมีโปรแกรมประยุกต์จำนวนไม่น้อยที่มีระบบสังเคราะห์เสียงเป็นส่วนสำคัญ เช่น โปรแกรมการสังเคราะห์ข้อความเสียง ซึ่งการเก็บสัมประสิทธิ์ที่เป็นคุณลักษณะสำคัญของสัญญาณเป็นจำนวนเพียงเล็กน้อยนับเป็นวิธีการที่มีประสิทธิภาพที่จะช่วยพัฒนาระบบในส่วนของการวิเคราะห์และสังเคราะห์เสียง ดังนั้น ในงานวิจัยนี้จึงได้ทำการศึกษาโครงสร้างหลักของคลื่นเสียงพยัญชนะและสระสำหรับการสังเคราะห์เสียงภาษาไทย โดยแบ่งออกเป็นสองกระบวนการใหญ่ ๆ คือ กระบวนการวิเคราะห์เสียง และกระบวนการสังเคราะห์เสียง ในส่วนของกระบวนการวิเคราะห์เสียง สัญญาณเสียงจะถูกแบ่งออกเป็นหน่วยเสียงส่วนต้น และหน่วยเสียงส่วนท้าย ซึ่งอยู่ในรูปของอัมพลิจูด จากนั้นจะทำการแยกองค์ประกอบของสัญญาณเสียงด้วยวิธีการแปลงฟูเรียร์ เพื่อคำนวณหาค่าพารามิเตอร์ที่เป็นคุณลักษณะสำคัญของสัญญาณเสียง อันได้แก่ ค่าแอมพลิจูด ค่าความถี่ และค่าเฟส และในส่วนของกระบวนการสังเคราะห์เสียงก็จะนำสัมประสิทธิ์ของพารามิเตอร์เหล่านั้นมาสังเคราะห์กลับเป็นสัญญาณเสียงด้วยแบบจำลองสัญญาณคลื่นไซน์ แล้วจึงนำสัญญาณไซน์หน่วยเสียงส่วนต้นและสัญญาณไซน์หน่วยเสียงส่วนท้ายมาเชื่อมต่อกันเป็นเสียงพยางค์ นอกจากนี้ในส่วนของการผันเสียงวรรณยุกต์จะใช้วิธีการปรับค่าพิชช์ เพื่อทำการผันเสียงวรรณยุกต์จากเสียงสามัญเป็นเสียงเอก เสียงโท เสียงตรี และเสียงจัตวา จากการทดลองพบว่าสามารถใช้จำนวนสัมประสิทธิ์ของพารามิเตอร์เพียง 15 เปอร์เซนต์เป็นตัวแทนของสัญญาณเสียงต้นแบบ มาทำการสังเคราะห์กลับเป็นสัญญาณเสียงที่ยอมรับได้

Speech synthesis system plays an important role in many applications, such as text-to-speech synthesizer. The effective and commonly used approach for implementing the system is analysis-and-synthesis scheme, since this scheme retains a few coefficients for coding the significant characteristics of signals. Therefore, in this thesis, we have investigated the infrastructure of consonants and vowels for synthesizing Thai voices by utilizing this approach. In this way, there are two important phases, analysis and synthesis phases. For the analysis phase, the speeches of consonants and vowels are acquired firstly and are segmented into demisyllable forms. Secondly, their characteristics are extracted by Fourier transform and the important parameters—amplitude, frequency, and phase—of the transform are coded. For the synthesis phase, the coefficients of the parameters analyzed by Fourier transform are reconstructed by means of sinusoidal model. A syllable voice is synthesized by concatenating the consonant and vowel demisyllables. Furthermore, in case of tonal voices, the pitch-period-control technique is applied to the normal tonal voices for generating 'Aek', 'To', 'Tree', and 'Juttawa' tones. By this approach, the experimental results show that the proposed method can synthesize the acceptable voices by using only 15% of the coefficients of Fourier transform.