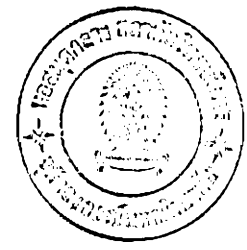


บทที่ 3

วิธีดำเนินการวิจัย

ในบทนี้จะกล่าวถึงรายละเอียดของขั้นตอนวิธีในการดำเนินการวิจัย ซึ่งประกอบไปด้วยรายละเอียดเกี่ยวกับการบันทึกเสียง การเก็บข้อมูลเสียงพูด การวิเคราะห์สัญญาณ การสังเคราะห์เสียง และโปรแกรมที่ใช้วิเคราะห์และสังเคราะห์เสียง

3.1 การเก็บข้อมูลเสียงพูด

การบันทึกเสียงพูดได้ทำการบันทึกเสียงที่ห้องปฏิบัติการวิจัยระบบเชิงเลข ตึกปฏิบัติการไมโครคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย ซึ่งได้ควบคุมสภาพแวดล้อมให้มีเสียงรบกวนน้อยที่สุด เสียงพูดที่บันทึกลงเครื่องคอมพิวเตอร์เก็บด้วยตัวอย่างขนาด 16 บิต อัตราการชักตัวอย่าง 22 กิโลเฮิรตซ์

การคัดเลือกผู้บอกภาษา

คุณสมบัติของผู้บอกภาษาที่ได้บันทึกเสียงเป็นเสียงต้นแบบในงานวิจัยนี้คือ

1. ผู้บอกภาษาเป็นเพศหญิง
2. เป็นผู้ที่มีการออกเสียงเป็นปกติและตรงตามหลักการออกเสียงพูดภาษาไทย
3. เป็นผู้ที่ใช้ภาษาไทยกรุงเทพเป็นภาษาพูดทั่วไป

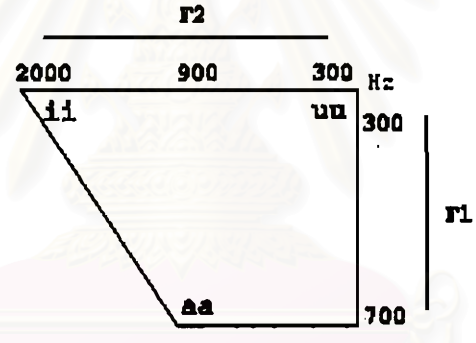
อุปกรณ์ที่ใช้ในการบันทึกเสียง

1. เครื่องคอมพิวเตอร์ Pentium 166 MMX หน่วยความจำ 32 MB Harddisk 1.2 GB และ Floppy Disk 1.44 MB
2. การ์ดเสียง Sound Blaster AWE 64 Value ของบริษัท Creative Technology
3. ไมโครโฟน AKG รุ่น D331BT
4. ระบบปฏิบัติการ Microsoft Windows 95 Thai Edition
5. โปรแกรม Goldwave เพื่อใช้ในการตัดต่อคลื่นเสียง
6. โปรแกรม Winsal เพื่อใช้ในการดูแผ่นภาพคลื่นเสียง

ในการสังเคราะห์เสียงเพื่อสร้างชุดคำสำหรับแบบทดสอบในการฟังได้ใช้สระ 3 หน่วยที่มีพยัญชนะ /s/ (ซ, ส) เป็นเสียงต้นพยางค์ สระทั้ง 3 หน่วยนี้คือ อา อี และอู แต่ละสระประกอบไปด้วยวรรณยุกต์จำนวนสระละ 5 วรรณยุกต์ รวมพยางค์ทั้งหมดที่ต้องบันทึกเสียงเป็นจำนวน 15 พยางค์ดังนี้

1. ซา ส่า ซ่า(ส้า) ซ้า สา จำนวน 5 พยางค์
2. ซู ซู่ ซู่(ซู้) ซู้ ซู จำนวน 5 พยางค์
3. ซี ซี่ ซี่(สี้) ซี้ สี จำนวน 5 พยางค์

ในงานวิจัยนี้ได้ทำการสังเคราะห์เสียงสระขึ้นมา 18 เสียง แต่เลือกมาทดสอบเพียงแค่ 3 สระนี้เท่านั้น เหตุที่เลือกสระ 3 ตัวนี้มาเป็นตัวแทนของทุกสระเนื่องจากว่าค่าความถี่ฟอร์แมนท์ที่ 1 ความถี่ฟอร์แมนท์ที่ 2 ซึ่งมีอิทธิพลมากต่อคุณภาพเสียงพูด ของสระทั้งสามตัวนี้แตกต่างกันมากที่สุด (วิสิทธิ์ ลีลาศิริวงศ์, 2535) ค่าความถี่ฟอร์แมนท์ที่ 2 ของสระอูและสระอีมีค่าแตกต่างกันมาก และค่าความถี่ฟอร์แมนท์ที่ 1 ของสระอากับสระอีและสระอู ก็ต่างกันมากเช่นกัน ดังรูปที่ 3.1



รูปที่ 3.1 แสดงแผนภาพสระแท้ (Cardinal Vowel)

จากรูปที่ 3.1 สระทั้ง 3 ตัวที่เลือกมามีความถี่ฟอร์แมนท์ที่ 1 และความถี่ฟอร์แมนท์ที่ 2 ครอบคลุมความเป็นไปได้ของสระที่จะเกิด ดังนั้นผลการทดสอบจากสระ 3 ตัวนี้น่าจะครอบคลุมถึงสระอื่นด้วย

3.2 การวิเคราะห์เพื่อเก็บฐานข้อมูลหน่วยเสียงอนุภาค

เสียงสระเป็นเสียงที่มีลักษณะเป็นรายคาบ รูปร่างจะคล้าย ๆ แต่ไม่เหมือนกันทั้งหมดรูปร่างจะแปรเปลี่ยนไปเรื่อย ๆ ตั้งแต่ต้นเสียงไปจนถึงท้ายเสียงแต่จะสมมติว่ามันเหมือนกัน เพื่อที่จะเก็บเพียงคาบเดียวของเสียงสระนี้เป็นหน่วยเสียงอนุภาค และเนื่องจากส่วนของเสียงสระนี้เป็นแกนของพยางค์จึงมีความยาวมากในพยางค์หนึ่ง ๆ ถ้าเราลดความยาวนี้เหลือเพียงหน่วยเสียงเพียงคาบเดียวจะทำให้ลดขนาดของฐานข้อมูลหน่วยเสียงลงได้มาก

หอสมุดกลาง จุฬาลงกรณ์มหาวิทยาลัย

ฐานข้อมูลที่จะต้องเก็บในงานวิจัยนี้ประกอบด้วย หน่วยเสียงพยัญชนะ หน่วยเสียงอนุภาคที่ตัดมาจากส่วนสระของพยางค์ ความถี่หลักมูลของเสียงสระ กรอบคลื่นแอมปลิจูดของเสียงสระ หรือมองอีกนัยหนึ่งคือจะสร้างส่วนของเสียงสระขึ้นมาใหม่ทั้งหมด โดยที่ส่วนพยัญชนะจะใช้หน่วยเสียงพยัญชนะเดิมที่ได้ตัดเก็บไว้

3.2.1 การเตรียมคลื่นเสียงก่อนนำมาวิเคราะห์

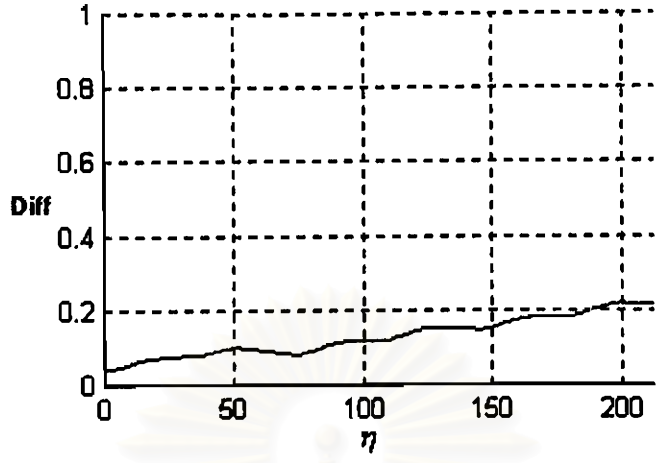
เสียงที่ได้บันทึกไว้ได้กำหนดให้บันทึกเป็นพยางค์ซึ่งแต่ละพยางค์จะมีวรรณยุกต์ที่แตกต่างกัน เช่นในเสียง ซา ก็จะมีการผันเป็น ส่า, ซ่า, ซ้า, สา ครบวรรณยุกต์ทั้ง 5 เสียง (ในที่นี้ยกตัวอย่างการวิเคราะห์เสียง ซา ส่า ซ่า ซ้า สา) จากนั้นจะทำการตัดแยกระหว่างพยัญชนะและสระออกจากกันโดยใช้โปรแกรม Goldwave รูปแบบที่แยกคือแยกแต่ละพยางค์ออกเป็นสองส่วนคือส่วนพยัญชนะและส่วนสระ เมื่อได้ตัดแยกระหว่างพยัญชนะและสระออกมาแล้วต่อไปจะนำส่วนของเสียงสระมาทำการวิเคราะห์หาค่าความถี่หลักมูล และกรอบคลื่นแอมปลิจูดต่อไป

ในการตัดแยกพยัญชนะและสระในพยางค์ออกจากกันจะพิจารณาจากแผ่นภาพคลื่นเสียงแบบช่วงการกรองกว้างประกอบด้วย เพื่อให้แน่ใจว่าได้แยกส่วนพยัญชนะและสระได้ถูกต้อง ในงานวิจัยนี้ได้ใช้โปรแกรม Winsal ของบริษัท Media Enterprise ในการแสดงแผ่นภาพคลื่นเสียง และได้ใช้โปรแกรม Goldwave ในการตัดต่อคลื่นเสียง

3.2.2 การหาค่าความถี่หลักมูล

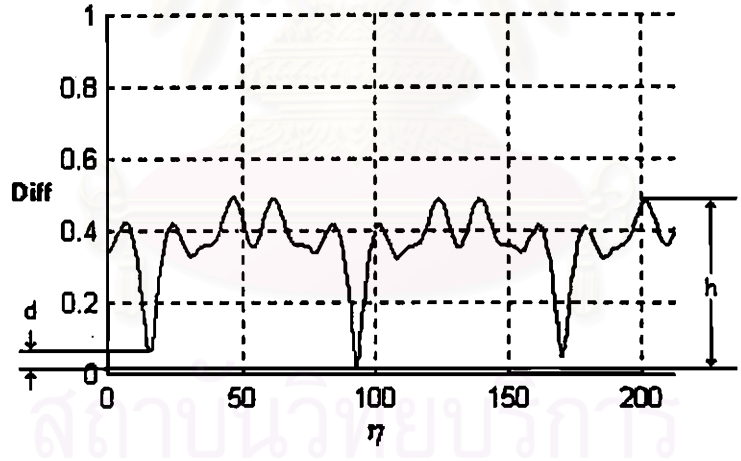
ความถี่หลักมูลคือความถี่หลักของเสียงพูด เพราะเสียงพูดเกิดจากการที่เส้นเสียงสั่นทำให้เกิดคลื่นผ่านไปยังช่องทางเดินเสียง และเกิดการกำทอนเป็นคลื่นความถี่สูงปนกับคลื่นเสียงจากเส้นเสียงผ่านปากออกมา คลื่นเสียงที่เกิดจากเส้นเสียงนั้นจะมีความถี่เท่ากับความถี่หลักมูลและมีความถี่กำทอนที่เกิดจากกล่องเสียงเนื่องจากความถี่หลักมูลด้วย ซึ่งเสียงที่ออกจากปากมาแล้วนั้นจะเป็นคลื่นเสียงที่มีความถี่ต่าง ๆ เหล่านี้ผสมกันอยู่ เราจะวิเคราะห์แยกเฉพาะความถี่หลักมูลออกมาเพื่อเก็บไว้เป็นส่วนหนึ่งของฐานข้อมูลหน่วยเสียง ในวิทยานิพนธ์นี้ได้ใช้กรรมวิธี AMDF ในการหาความถี่หลักมูล

กรรมวิธี AMDF สามารถใช้ตัดสินว่าสัญญาณเสียงที่นำมาวิเคราะห์นั้นเป็นเสียงก้องหรือไม่ก้องได้ คือสามารถใช้ตัดสินว่าช่วงไหนของสัญญาณเป็นพยัญชนะหรือเป็นสระ ถ้าสัญญาณเสียงที่นำมาวิเคราะห์เป็นเสียงก้องกรรมวิธีนี้จะบอกค่าความถี่หลักมูลของสัญญาณนี้ได้ ถ้าสัญญาณที่วิเคราะห์เป็นเสียงไม่ก้องผลลัพธ์ที่ได้จะเป็นกราฟที่ไม่มีจุดต่ำสุดที่เป็นระเบียบ หรืออาจได้ผลลัพธ์เป็นเส้นเรียบดังรูปที่ 3.2



รูปที่ 3.2 แสดงผลลัพธ์จากกรรมวิธี AMDF กับเสียงไม่ก้อง

ถ้าสัญญาณที่นำมาวิเคราะห์เป็นเสียงก้อง จะได้กราฟที่มีจุดต่ำสุดที่เรียงกันอย่างเป็นระเบียบและมีขนาดของความแตกต่างค่อนข้างสูงดังรูปที่ 3.3

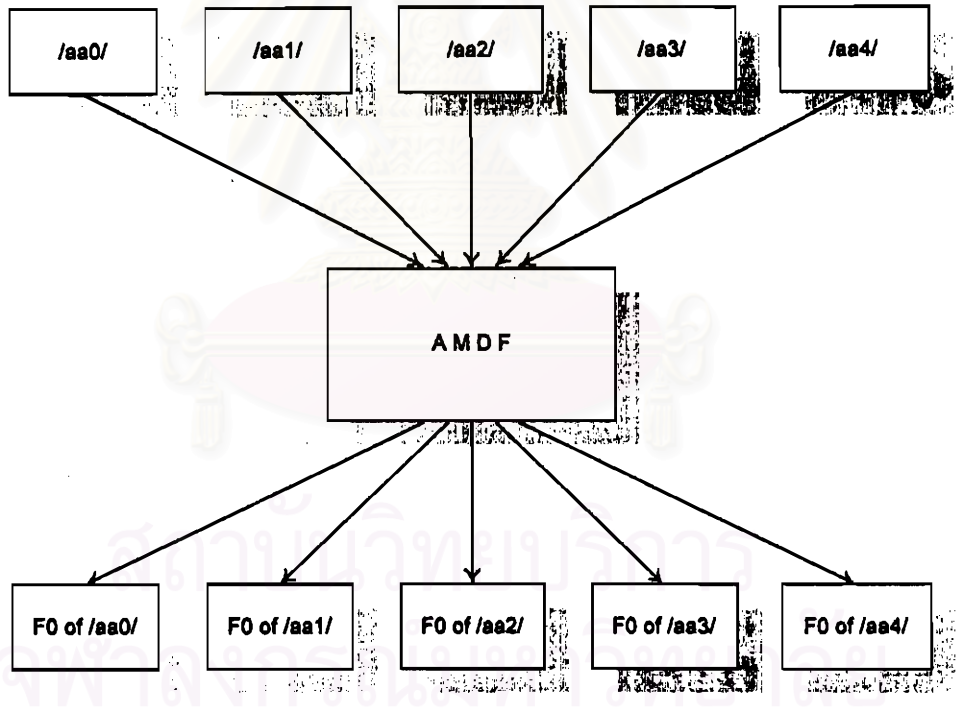


รูปที่ 3.3 แสดงผลลัพธ์จากกรรมวิธี AMDF กับเสียงก้อง

ที่จุดต่ำสุดแต่ละค่าเป็นตัวบอกว่ารูปร่างของสัญญาณที่นำมาวิเคราะห์คล้ายกันมาก หมายความว่าเกิดการครบคาบของสัญญาณที่ตัวอย่างเสียงถูกชักมา ดังนั้นระยะห่างระหว่างจุดต่ำสุดสองค่านี้ก็คือความยาวของคาบสัญญาณเป็นจำนวนสัญญาณที่ถูกชักตัวอย่างมา ถ้าเอาค่านี้ไปหารความถี่ในการชักตัวอย่างก็จะได้ค่าความถี่หลักมูล ถ้าหาค่าความถี่หลักมูลจากระยะห่างระหว่างจุดต่ำสุดสองค่าอาจเป็นการยุ่งยาก แต่สามารถพิจารณาจากจุดต่ำสุดจุดแรกได้เพราะสะดวกต่อการหามากกว่า ในบางครั้งจุดต่ำจุดแรกที่ได้อาจมีค่าต่ำไม่เท่ากับจุดต่ำสุดจุดที่สอง แต่ถ้าเกิดกรณีเช่นนี้จะ

ถือว่าจะเอาจุดต่ำสุดจุดแรกมาใช้ ดังนั้นจึงมีการกำหนดค่า d ไว้ที่ค่า ๆ หนึ่งซึ่งถ้าค่า d มีค่าไม่เกินค่าที่ตั้งไว้ จะถือว่าจุดต่ำสุดจุดแรกคือจุดที่ต้องการ แต่ถ้ามากกว่าแสดงว่าจุดต่ำสุดจุดถัดไปคือจุดที่ต้องการ หรือในบางกรณีเช่นในกรณีของพยัญชนะเสียงก้องกรรมวิธี AMDF นี้จะให้ผลลัพธ์ที่เหมือนกับผลที่ได้จากสัญญาณเสียงก้อง แต่ความแตกต่างระหว่างค่ามากที่สุดและค่าน้อยที่สุดในกราฟจะต่างกันไม่มาก คือค่า h จะน้อย จึงมีการกำหนดค่า h ไว้เป็นค่าเปรียบเทียบด้วยว่า ถ้าค่า h มีค่าน้อยกว่าค่าที่ตั้งไว้จะถือว่าสัญญาณเสียงดังกล่าวเป็นสัญญาณเสียงไม่ก้อง ถ้าค่า h มากกว่าค่าที่ตั้งไว้สัญญาณดังกล่าวเป็นสัญญาณเสียงก้อง

ในการหาความถี่หลักมูลนี้ จะนำเสียงสระแต่ละวรรณยุกต์ที่ได้ตัดเตรียมไว้มาทำกรรมวิธี AMDF ในที่นี้คือ /aa0/, /aa1/, /aa2/, /aa3/ และ /aa4/ เสร็จแล้วจะให้ทางเดินของความถี่หลักมูลออกมา 5 ชุด และจะนำหน่วยเสียงสระไปแยกเอาหน่วยเสียงอนุภาคออกมาเป็นขั้นตอนต่อไป



รูปที่ 3.4 แสดงค่าความถี่หลักมูลที่ได้มาจากกรรมวิธี AMDF

เนื่องจากค่าความถี่หลักมูลมีการเปลี่ยนแปลงไปตามเวลาไม่ได้คงที่ตลอดทั้งพยางค์ ดังนั้นจึงต้องแบ่งสัญญาณเสียงพูดที่จะนำมาวิเคราะห์ด้วยกรรมวิธี AMDF เป็นกรอบเสียงพูดหลาย ๆ กรอบ ความยาวของกรอบเสียงพูดจะต้องครอบคลุมมากกว่าสองคาบของสัญญาณ ในหนึ่งกรอบ

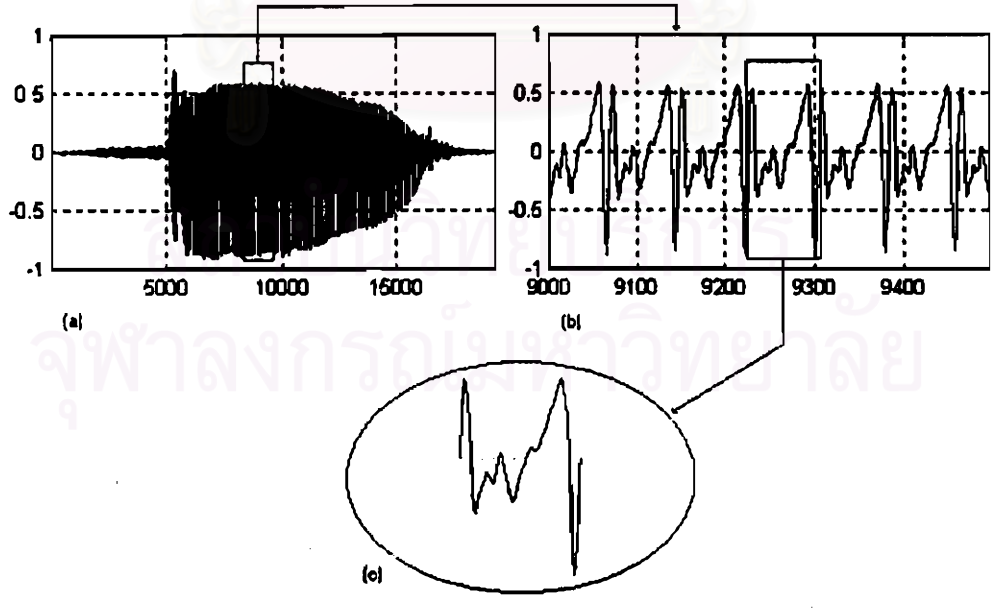
สัญญาณเสียงพูดที่นำมาวิเคราะห์ด้วยกรรมวิธี AMDF จะได้ค่าความถี่หลักมูลออกมาหนึ่งค่า และได้ให้กรอบเสียงพูดเหลื่อมทับกัน 75% เพื่อให้ความละเอียดทางเวลาเพิ่มมากขึ้น

3.2.3 การเลือกหน่วยเสียงอนุภาค

จากส่วนสระที่ได้ตัดเตรียมไว้ในขั้นตอนก่อนหน้านี้ จะนำมาตัดหน่วยเสียงอนุภาคออกมาเก็บเป็นฐานข้อมูล ในส่วนของเสียงสระถ้านำมาขยายเพื่อดูในช่วงสั้น ๆ จะพบว่ามันจะรูปคลื่นเป็นรายคาบซึ่งมีรูปร่างคล้าย ๆ กันเรียงกันไปจนสุดเสียง รูปร่างของรูปคลื่นจะค่อย ๆ มีการเปลี่ยนแปลงไปทีละน้อยถ้าเปรียบเทียบจากช่วงต้นเสียงไปจนถึงช่วงท้ายเสียง แต่จะเปลี่ยนแปลงไปไม่มากจนผิดรูปร่าง

ในขั้นต้นนี้ได้ทำการเลือกหน่วยเสียงอนุภาคมาจากบริเวณกึ่งกลางของเสียงสระ เนื่องจากมันมีพลังงานค่อนข้างมากและน่าจะเป็นตัวแทนที่ดี มีการเปลี่ยนแปลงรูปร่างไม่มากเมื่อเปรียบเทียบกับหน่วยเสียงอนุภาคในช่วงต้นและหน่วยเสียงอนุภาคในช่วงท้าย ถ้าตัดหน่วยเสียงอนุภาคจากช่วงต้นมาเป็นตัวต้นแบบ เมื่อเทียบกับหน่วยเสียงอนุภาคช่วงท้ายอาจมีการเปลี่ยนแปลงไปมาก เมื่อนำมาเป็นหน่วยเสียงต้นแบบเพื่อสังเคราะห์เสียงคุณภาพที่ได้อาจด้อยลงไป

ในรูปที่ 3.5 ข้างล่างนี้แสดงถึงการแยกตัดหน่วยเสียงอนุภาคจากบริเวณกลางพยางค์ ในการตัดได้ทำการตัดหน่วยเสียงอนุภาคบริเวณที่ตัดผ่านจุดศูนย์ เพื่อให้การนำมาเชื่อมต่อภายหลังต่อได้สมบูรณ์ขึ้น



รูปที่ 3.5 แสดงการเลือกตัดหน่วยเสียงอนุภาค

3.2.4 การหากรอบคลื่นแอมปลิจูด

กรอบคลื่นแอมปลิจูดจะถูกใช้ในการเปลี่ยนแปลงให้แอมปลิจูดของพยางค์เป็นไปตามเสียงต้นแบบเพื่อให้เสียงสังเคราะห์มีพลังงานเหมือนกับเสียงต้นแบบ เพราะแต่ละสระและวรรณยุกต์ก็จะมีคุณลักษณะของพลังงานเสียงเฉพาะตัว เช่นในวรรณยุกต์จัตวาจะมีความถี่หลักมูลในช่วงท้ายพยางค์เพิ่มขึ้นค่อนข้างสูงและปัจจัยหนึ่งที่ทำให้ความถี่หลักมูลมีค่าสูงคือ แรงลมจากช่องท้องที่ผ่านเส้นเสียงมีแรงมากทำให้เส้นเสียงสั่นเร็วเกิดเป็นความถี่สูงขึ้นมา กล่าวคือวรรณยุกต์จัตวามีพลังงานเพิ่มขึ้นในช่วงท้ายพยางค์ซึ่งในวรรณยุกต์ตัวอื่นจะไม่มีคุณลักษณะเช่นนี้

ในการหากรอบคลื่นแอมปลิจูดนี้ทำโดยแบ่งเสียงพูดออกเป็นกรอบ ในแต่ละกรอบเสียงพูดที่เลื่อนไปจะหาค่ามากที่สุดและค่าน้อยที่สุดของแอมปลิจูดเก็บไว้เฉพาะแต่ละกรอบเสียงพูด เมื่อทำเช่นนี้จนครบความยาวเสียงก็จะได้กรอบคลื่นแอมปลิจูดซึ่งเป็นค่าประมาณกรอบคลื่นจริง ๆ ออกมา ในขั้นตอนนี้ได้เพิ่มความละเอียดของการหาโดยการให้กรอบเสียงพูดที่นำมาวิเคราะห์เหลื่อมซ้อนกันด้วย

สรุปขั้นตอนการวิเคราะห์เพื่อหาองค์ประกอบของเสียงเก็บไว้เป็นฐานข้อมูล คือบันทึกเสียงพยางค์ต้นแบบและจากพยางค์ต้นแบบจะนำมาตัดแยกส่วนพยัญชนะและสระ ส่วนของสระจะนำมาหาความถี่หลักมูล ตัดหน่วยเสียงอนุภาค และหากรอบคลื่นแอมปลิจูด เก็บไว้เป็นฐานข้อมูลพร้อมส่วนพยัญชนะ เพื่อจะนำมาทำการสังเคราะห์เสียงต่อไป

3.3 การสังเคราะห์เสียงจากหน่วยเสียงอนุภาค

จากขั้นตอนการวิเคราะห์หาองค์ประกอบเพื่อเก็บเป็นฐานข้อมูลหน่วยเสียง จะได้หน่วยเสียงที่เก็บไว้เป็นฐานข้อมูลคือ พยัญชนะ ค่าความถี่หลักมูล หน่วยเสียงอนุภาค และกรอบคลื่นแอมปลิจูด ในขั้นตอนการสังเคราะห์เสียงนี้ จะนำองค์ประกอบที่ได้วิเคราะห์เก็บไว้นั้นมาสังเคราะห์กลับเป็นเสียงพยางค์ตามที่ต้องการ โดยนำหน่วยเสียงอนุภาคของสระที่ต้องการมาเชื่อมต่อกันให้ได้ความยาวตามที่ต้องการ ในขณะที่นำหน่วยเสียงอนุภาคมาเชื่อมต่อกันจะมีการเปลี่ยนค่าความถี่หลักมูลของหน่วยเสียงอนุภาคพร้อมกับเปลี่ยนค่าแอมปลิจูดด้วยค่าความถี่หลักมูลและกรอบคลื่นแอมปลิจูดที่เก็บเอาไว้

3.3.1 การเปลี่ยนค่าความถี่หลักมูล

การที่จะสังเคราะห์เสียงกลับให้มีวรรณยุกต์ต่าง ๆ ตามที่ต้องการ จะต้องเปลี่ยนแปลงความถี่หลักมูลของหน่วยเสียงอนุภาคให้เป็นไปตามค่าความถี่หลักมูลของวรรณยุกต์นั้น ๆ ความถี่หลักมูลของสัญญาณคือคาบเวลาตั้งแต่สัญญาณเริ่มเกิดจนกระทั่งมาเกิดซ้ำอีกครั้ง มีจำนวนตัวอย่างที่ถูกชักมากี่ตัวในหนึ่งคาบซึ่งคาบเวลาสามารถหาได้จากจำนวนตัวอย่างที่ถูกชักมาในหนึ่งคาบเมื่อนำมาหา

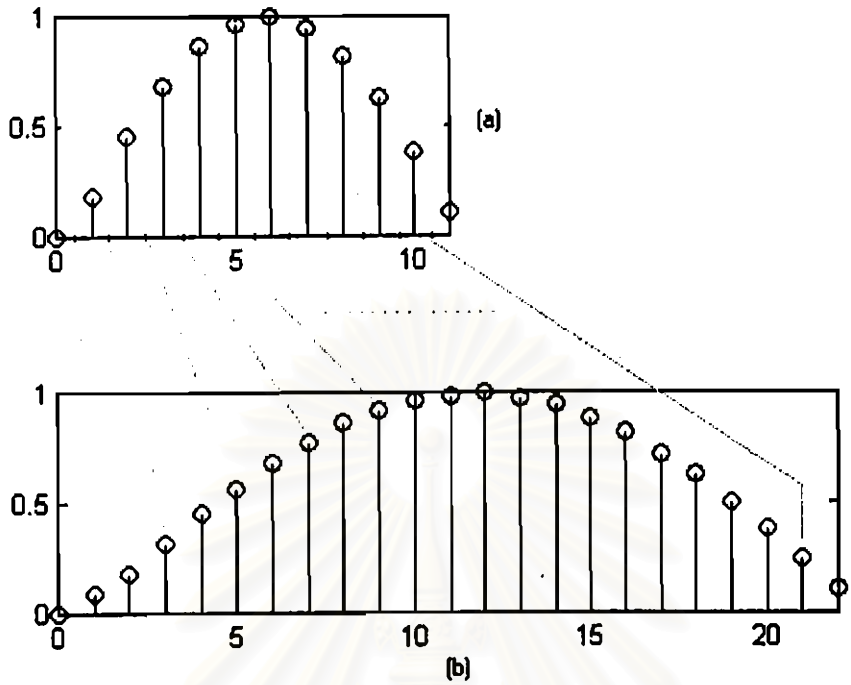
ส่วนกลับของคาบเวลานี้ก็จะได้ค่าความถี่หลักมูล ซึ่งคาบเวลานี้ถูกบังคับด้วยอัตราชักตัวอย่างคือค่าความถี่ Fs ตัวอย่างที่ถูกชักมาแต่ละตัวจะมีคาบเวลาห่างกันเท่ากับ $\frac{1}{Fs}$ วินาที ถ้ามีตัวอย่างจำนวน n ตัวในหนึ่งคาบ สัญญาณนี้จะมีความถี่หลักมูลเท่ากับ $\frac{Fs}{n}$ เฮิรตซ์ เราสามารถที่จะเ พิ่มหรือลดจำนวนตัวอย่างที่ถูกชักมาได้ตามสมการ

$$y[n] = x[Mn] \tag{3.1}$$

ซึ่งถ้า $M < 1$ จะเป็นการลดค่าความถี่หลักมูล และ $M > 1$ จะเป็นการเพิ่มค่าความถี่หลักมูล โดยที่ค่า M ไม่จำเป็นที่ต้องเป็นเลขจำนวนเต็ม เพื่อที่จะเปลี่ยนให้สัญญาณมีค่าคาบใหม่ได้โดยที่รูปร่างของสัญญาณยังคงเหมือนเดิมซึ่งเสมือนว่าได้เปลี่ยนค่าความถี่หลักมูลของสัญญาณนั้น ที่ $M < 1$ สัญญาณ $y[n]$ มีจะระยะคาบยาวขึ้นทำให้มีค่าความถี่หลักมูลลดลง ที่ $M > 1$ จะทำให้สัญญาณ $y[n]$ มีระยะคาบสั้นลงซึ่งจะทำให้มีค่าความถี่หลักมูลเพิ่มขึ้น

จากสมการข้างบนจะเห็นว่าในกรณีที่ M ไม่ใช่จำนวนเต็ม ค่า Mn ก็จะไม่ใช่จำนวนเต็มด้วยซึ่งไม่ตกตรงตำแหน่งของตัวอย่างที่มีอยู่ ดังนั้นจึงจำเป็นต้องประมาณค่าตัวอย่างนี้ขึ้นมาจากค่าตัวอย่างที่อยู่ใกล้เคียงเพื่อทำให้จำนวนตัวอย่างมีจำนวนเท่ากับคาบเวลาที่เพิ่มขึ้นหรือลดลง ซึ่งในวิทยานิพนธ์นี้ได้ใช้กรรมวิธีการประมาณค่าในช่วงแบบเชิงเส้นมาทำการประมาณค่าตัวอย่างดังกล่าว ซึ่งกรรมวิธีนี้เพียงพอที่จะนำมาใช้ได้เนื่องจากว่าได้ทำการชักตัวอย่างสัญญาณเสียงพูดมาด้วยความถี่ที่ค่อนข้างสูง ในรูปที่ 3.6 เป็นแนวคิดที่จะประมาณค่าตัวอย่างขึ้นมาใหม่เพื่อให้คาบเวลายาวนานขึ้นโดยรูปร่างของคลื่นยังเหมือนเดิม เพื่อเป็นการเปลี่ยนค่าความถี่หลักมูลของสัญญาณ

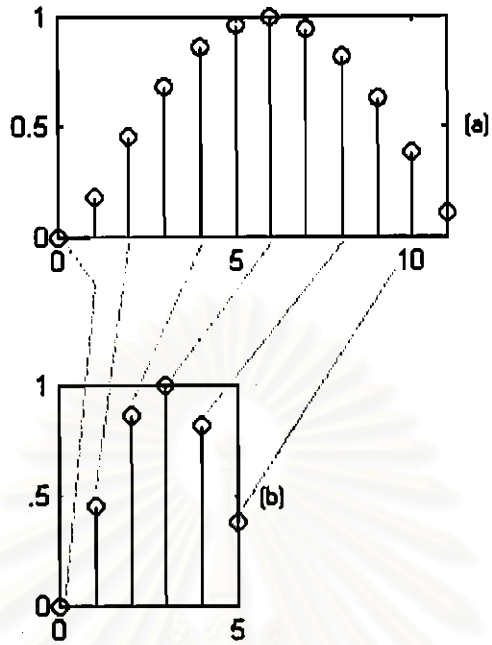
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.6 แสดงการลดค่าความถี่หลักมูลของสัญญาณ

จากรูปที่ 3.6 (a) เป็นสัญญาณต้นแบบ ต้องการที่จะลดความถี่หลักมูลลงไปครึ่งหนึ่ง คือจะต้องเพิ่มจำนวนตัวอย่างขึ้นมาเป็นจำนวน 2 เท่า ด้วยค่า $M = 0.5$ เพื่อจะทำให้ได้ระยะเวลาในการครบคาบมากขึ้น 2 เท่าเพื่อเป็นผลให้ความถี่หลักมูลลดลงครึ่งหนึ่ง เราจะต้องหาตัวอย่างตัวใหม่มาแทรกระหว่างตัวอย่างที่มีอยู่แล้วในรูปที่ 3.6 (a) เช่นระหว่างตัวอย่างที่ 0 และ 1 ตัวอย่างที่ 1 และ 2 เช่นนี้เรียงลำดับไป เช่นในระหว่างตัวอย่างที่ 1 และ 2 เรารู้ค่าของเฉพาะสองตัว ไม่รู้ค่าตรงกลางที่ต้องการจะหาคือแอมปลิจูดของตัวอย่างลำดับที่ 1.5 แต่เราสามารถหาได้โดยใช้ค่าของสองตัวที่รู้แล้วด้วยการประมาณค่าในช่วงเชิงเส้น จะเหมือนว่าลากเส้นตรงระหว่างจุดสองจุดที่รู้ค่าคือระหว่างตัวอย่างที่ 1 และ 2 จากนั้นจะสามารถคำนวณหาค่าบนแกน y ได้เมื่อรู้ค่าบนแกน x ที่อยู่บนเส้นตรงนี้ ในรูปที่ 3.6 (b) ที่เห็นเส้นตรงลากจากรูปที่ 3.6 (a) ลงมาคือค่าตัวอย่างใหม่ที่ได้จากการประมาณค่า ซึ่งจะเห็นว่ารูปคลื่นยังคงมีรูปร่างเหมือนเดิมในขณะที่คาบของสัญญาณได้เพิ่มขึ้น

ในรูปที่ 3.7 เป็นการเพิ่มความถี่หลักมูลของสัญญาณเป็น 2 เท่าจากค่าความถี่หลักมูลของสัญญาณเดิมโดยใช้ค่า $M = 2$ บังเอิญว่าในตัวอย่างที่ยกมาเป็นการเพิ่มความถี่หลักมูลเป็น 2 เท่า ทำให้ตำแหน่งที่จะเกิดค่าใหม่ตกลงบนตัวอย่างที่อยู่ในลำดับคู่ ดังนั้นจะเห็นว่าในรูปที่ 3.7 (b) เป็นการดึงเอาตัวอย่างในลำดับคู่มาจากรูปที่ 3.7 (a) แต่ในบางกรณีที่การเพิ่มความถี่หลักมูลไม่ได้เป็นจำนวนเต็ม ตัวอย่างที่จะเกิดขึ้นจะตกอยู่ในช่วงระหว่างตัวอย่างสองตัว ซึ่งก็สามารถที่จะประมาณค่าได้โดยใช้การประมาณค่าในช่วงเชิงเส้นเช่นกัน



รูปที่ 3.7 แสดงการเพิ่มค่าความถี่หลักมูลของสัญญาณ

ค่าความถี่หลักมูลที่ได้จากการวิเคราะห์ จะมีจำนวนเท่ากับจำนวนกรอบเสียงพูดที่ได้แบ่งไว้ ซึ่งในขั้นตอนการสังเคราะห์จะต้องนำค่าความถี่หลักมูลนั้นมาปรับเปลี่ยนกับหน่วยเสียงอนุภาค ซึ่งมีจำนวนหนึ่งที่อาจจะไม่เท่ากับค่าความถี่หลักมูลที่ได้เก็บไว้ เนื่องจากจำนวนหน่วยเสียงอนุภาคที่จะนำมาต่อกันนั้นขึ้นอยู่กับความยาวของเสียงที่ต้องการสังเคราะห์ ดังนั้นจึงจะต้องมีการเพิ่มหรือลดจำนวนของค่าความถี่หลักมูลที่ได้เก็บไว้ก่อนนั้นให้มีจำนวนเท่ากับจำนวนหน่วยเสียงอนุภาคที่จะนำมาต่อกันเสียก่อนโดยใช้การประมาณค่าในช่วงแบบเชิงเส้นเช่นกันกับค่าความถี่หลักมูลที่เก็บไว้ เพื่อให้ได้จำนวนของค่าความถี่หลักมูลใหม่ตามที่ต้องการ

3.3.2 การเปลี่ยนค่าแอมปลิจูด

ค่าแอมปลิจูดที่เก็บไว้ในการวิเคราะห์จะมีจำนวนเท่ากับจำนวนกรอบของเสียงพูดที่แบ่งไว้เช่นกันซึ่งอาจจะมีจำนวนไม่เท่ากับจำนวนหน่วยเสียงอนุภาคที่จะต้องนำมาต่อกัน ดังนั้นจึงต้องมีการประมาณค่าขึ้นมาใหม่เพื่อให้จำนวนของค่าแอมปลิจูดมีจำนวนเท่ากับจำนวนของหน่วยเสียงอนุภาคเสียก่อน เนื่องจากว่ากรอบคลื่นแอมปลิจูดที่หาเตรียมไว้นั้นมีทั้งขอบช่วงบวกและช่วงลบ การที่จะนำมาเปลี่ยนแปลงค่าแอมปลิจูดของหน่วยเสียงอนุภาคจะต้องตรวจก่อนว่าตัวอย่างไหนของหน่วยเสียงอนุภาคที่อยู่ในช่วงบวกจะนำมาคูณกับอัตราส่วนต่างระหว่างกรอบคลื่นแอมปลิจูดช่วงบวกกับค่า

แอมปลิจูดของตัวอย่างนั้น และตัวอย่างไหนที่อยู่ในช่วงลบจะนำมาคูณกับอัตราส่วนต่างระหว่างกรอบคลื่นแอมปลิจูดช่วงลบกับค่าแอมปลิจูดของตัวอย่างนั้น

3.3.4 ความยาวของพยางค์ที่สังเคราะห์

ปิยฉัตร ปานโรจน์ (2534) ได้ทำการวิจัยพบว่าลักษณะเชิงกลสัทศาสตร์ของวรรณยุกต์ภาษาไทยมีการแปรกลุ่มตามอายุ ค่าความยาวของพยางค์ที่ถูกเปล่งออกมาจะมีความยาวที่แตกต่างกันไป ดังตารางที่ 3.1 และตารางที่ 3.2 สามารถที่จะนำค่าระยะเวลานี้มาเป็นตัวกำหนดความยาวเสียงที่จะสังเคราะห์ขึ้นมาใหม่ เพื่อให้เสียงสังเคราะห์มีความยาวของพยางค์เท่ากับเสียงพูดธรรมชาติ

วรรณยุกต์	โครงสร้างพยางค์	ระยะเวลา (ms)		
		50-60 ปี	30-40 ปี	10-20 ปี
เอก	พยางค์เป็นสระยาว	415	414	445
	พยางค์เป็นสระสั้น	366	359	392
	พยางค์ตายสระยาว	297	303	306
	พยางค์ตายสระสั้น	130	149	154
สามัญ	พยางค์เป็นสระยาว	438	421	459
	พยางค์เป็นสระสั้น	377	385	422
	พยางค์เป็นสระยาว	436	401	466
ตรี	พยางค์ตายสระสั้น	358	349	389
	พยางค์สระสั้น	148	148	196

ตารางที่ 3.1 การแปรค่าระยะเวลาของวรรณยุกต์คงระดับ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วรรณยุกต์	โครงสร้างพยางค์	ระยะเวลา (ms)		
		50-60 ปี	30-40 ปี	10-20 ปี
โท	พยางค์เป็นสระยาว	402	395	417
	พยางค์เป็นสระสั้น	329	331	378
	พยางค์ตายสระยาว	305	293	363
จัตวา	พยางค์เป็นสระยาว	423	422	426
	พยางค์เป็นสระสั้น	352	357	376

ตารางที่ 3.2 การแปรค่าระยะเวลาของวรรณยุกต์เปลี่ยนระดับ

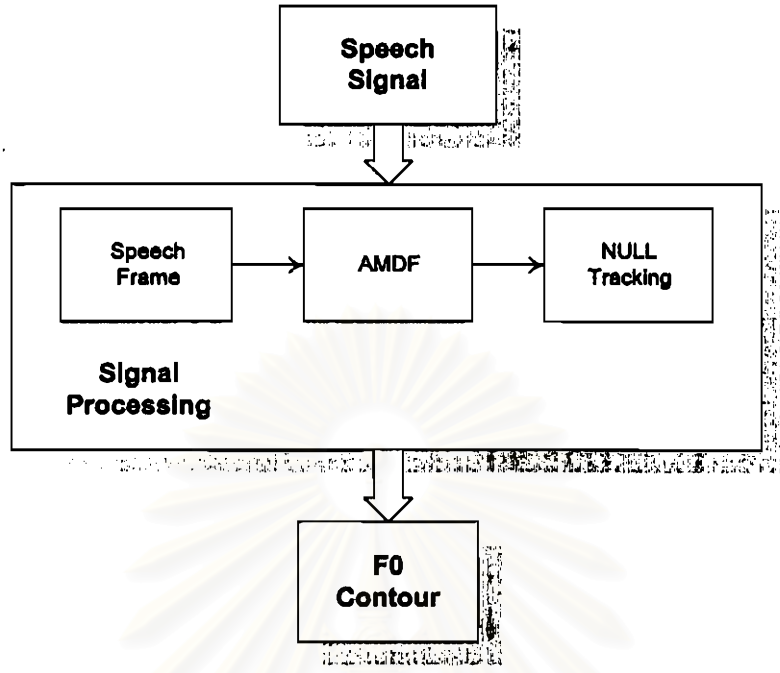
3.3.3 การเชื่อมต่อหน่วยเสียงอนุภาค

เมื่อได้ปรับเปลี่ยนความถี่หลักมูลและกรอบคลื่นแอมปลิจูดของหน่วยเสียงอนุภาคให้เป็นไปตามสระและวรรณยุกต์ที่ต้องการซึ่งใช้ค่าความถี่หลักมูลและค่ากรอบคลื่นแอมปลิจูดที่ได้วิเคราะห์จากเสียงจริงเก็บไว้แล้ว จะนำหน่วยเสียงอนุภาคที่ได้ถูกปรับเปลี่ยนความถี่หลักมูลและแอมปลิจูดแล้วนี้มาเชื่อมต่อกัน และสุดท้ายนำไปเชื่อมต่อกับพยัญชนะที่ต้องการซึ่งในการเชื่อมต่อจะใช้วิธีเชื่อมต่อตามวิธีของสุดาพร ลักษณียนาวิน (2535)

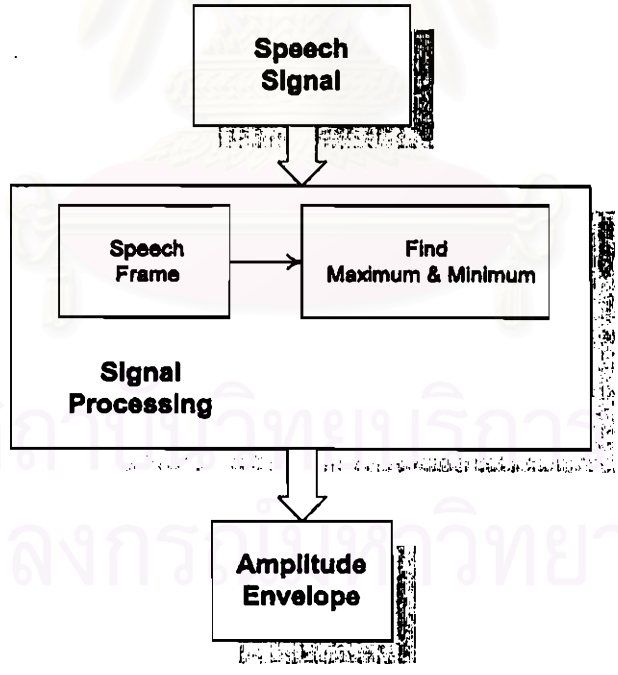
วิทยานิพนธ์นี้ได้เขียนโปรแกรมวิเคราะห์และสังเคราะห์เสียงด้วยโปรแกรม Matlab เนื่องจากว่ารวดเร็วในการพัฒนาระบบต้นแบบและมีฟังก์ชั่นต่าง ๆ เตรียมไว้ให้ใช้ โปรแกรมทั้งหมดจะมีทั้งโปรแกรมสำหรับวิเคราะห์องค์ประกอบและโปรแกรมสังเคราะห์เสียง ซึ่งประกอบไปด้วย

1. AMDF สำหรับหาค่าความถี่หลักมูล
2. Find_envelope สำหรับหากรอบคลื่นแอมปลิจูด
3. MP_SYNT สำหรับสังเคราะห์เสียง

ซึ่งแต่ละโปรแกรมมีแผนผังการทำงานดังนี้

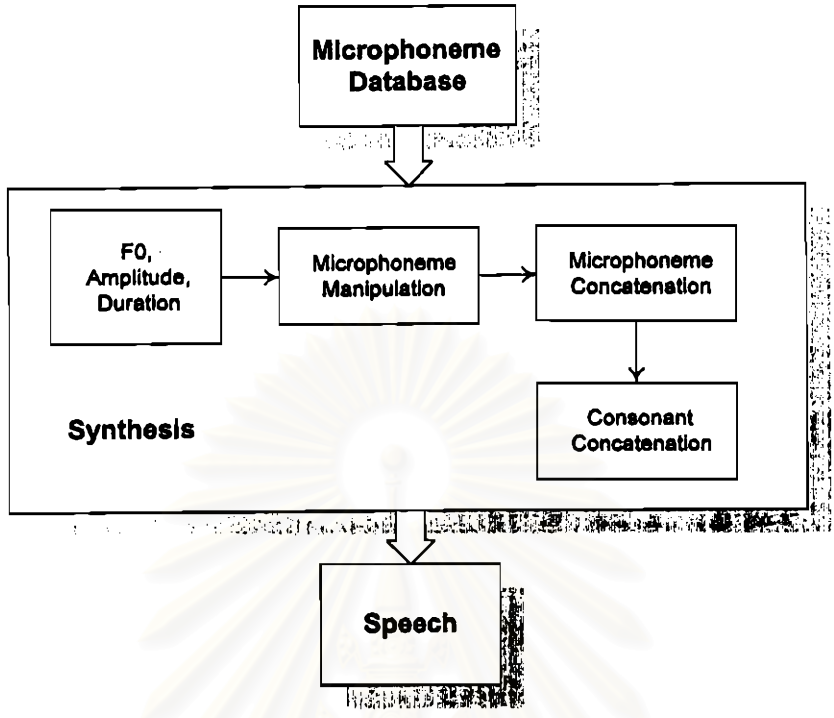


รูปที่ 3.8 แผนผังการทำงานของโปรแกรม AMDF



รูปที่ 3.9 แผนผังการทำงานของโปรแกรม Find_envelope

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.10 แผนผังการทำงานของโปรแกรม MP_SYNT

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย