

ต้นแบบโปรแกรมคอมพิวเตอร์ แปลงเอกสารภาษาไทยเก็บเป็นแฟ้มข้อมูล



▲ ภาพแสดงหน้าปกผลิตภัณฑ์ โปรแกรมอ่านไทย เวอร์ชัน 2.0

โปรแกรมคอมพิวเตอร์แปลงเอกสารภาษาไทยเก็บเป็นแฟ้มข้อความนี้สามารถนำไปเชื่อมต่อกับโปรแกรมประยุกต์อื่นๆ ที่ต้องการคุณสมบัติการแปลงเอกสารภาษาไทยเป็นข้อความได้ รวมทั้งปรับปรุงให้ใช้งานกับลักษณะงานอื่นๆ ที่ใกล้เคียงกันได้ไม่ยากนัก เช่น การรู้จำเฉพาะตัวเลขในเอกสาร หรือรู้จำตัวอักษรจากรูปภาพต่างๆ เช่น รูปภาพป้ายทะเบียนรถ เป็นต้น

ลักษณะการดำเนินการ

การพัฒนาเทคโนโลยีไอซีอาร์ประกอบด้วย 3 ขั้นตอน ได้แก่

1. ขั้นตอนการประมวลผลภาพลักษณะก่อนการรู้จำ (Pre-processing)
2. ขั้นตอนการรู้จำตัวอักษร (Character Recognition)
3. ขั้นตอนการประมวลผลหลังการรู้จำ (Post-processing)

คุณสมบัติ

โปรแกรมที่ได้สามารถผลิตเป็นซอฟต์แวร์เชิงพาณิชย์ในนาม “อ่านไทย เวอร์ชัน 2.0” โดยมีคุณสมบัติดังนี้

- รับภาพจากสแกนเนอร์ได้ โดยผ่านทางทเวิน (TWAIN)
- เปิดเพิ่มข้อมูลภาพได้ทั้งภาพสี ขาวดำ และภาพสีเทา ได้หลายรูปแบบ เช่น bmp, tif, gif, pcx, jpg โดยไม่จำกัดขนาดของแฟ้มข้อมูล และไม่จำกัดจำนวนแฟ้มข้อมูลที่เปิด
- ปรับแต่งภาพได้ตามต้องการ แปลงภาพจากภาพสีเป็นภาพสีเทา ภาพสีเทาเป็นภาพขาวดำ หรือภาพขาวดำกลับเป็นภาพสีเทา รวมทั้งการสลับสีจากขาวเป็นดำ ดำเป็นขาว การย่อหรือขยายภาพ การหมุน 90 องศา 180 องศา 270 องศา รวมทั้งการปรับหน้าที่เอียงให้ตรงแบบอัตโนมัติ หรือให้ผู้ใช้งานใส่มุมที่ต้องการปรับเอง
- สามารถแปลงเอกสารได้ทั้งภาษาไทยและภาษาอังกฤษ ทั้งตัวอักษรขนาดเล็กและขนาดใหญ่ มีความถูกต้องประมาณ 95 % รู้จักฟอนต์ทั้งที่อยู่บน Window และฟอนต์ที่อยู่บนเครื่อง Macintosh ซึ่งเป็นฟอนต์ของโรงพิมพ์ต่างๆ
- สามารถสั่งการทั้งแบบที่ละแฟ้มข้อมูล (online mode) และแบบสั่งการทีละหลายๆแฟ้มข้อมูลพร้อมกัน (batch mode)
- มีการตรวจคำสะกดโดยจะแสดงการไฮไลต์คำผิดให้อัตโนมัติและจะแสดงคำใกล้เคียงให้เลือกเพื่อแทนที่คำที่ผิด หรือจะแก้ไขคำผิดโดยตรงก็ได้
- แก้ไขหรือเพิ่มเติมข้อความได้ทันที และสามารถตัดหรือคัดลอกไปใส่เวิร์ดโปรเซสเซอร์อื่นๆได้ นอกจากนี้ยังสามารถคัดลอกภาพในเอกสารมาใส่ในข้อความได้อีกด้วย เนื่องจากสามารถเก็บเป็นแฟ้มข้อความรูปแบบ rich text format (rtf) ได้ด้วย

คุณสมบัติของระบบที่ต้องการ

โปรแกรมอ่านไทย เวอร์ชัน 2.0 ต้องการระบบฮาร์ดแวร์และซอฟต์แวร์ ต่อไปนี้

- เครื่องคอมพิวเตอร์ที่มีระบบปฏิบัติการ Window 95 หรือสูงกว่า ที่มีระบบภาษาไทย
- หน่วยความจำอย่างต่ำ 16 เมกะไบต์
- เนื้อที่ว่างในฮาร์ดดิสก์ประมาณ 10 เมกะไบต์
- ความละเอียดจอภาพ 600x800 หรือสูงกว่า

ประโยชน์ที่ได้รับ

1. ประหยัดทรัพยากรแรงงานและทรัพยากรเวลาของมนุษย์ ด้านงานพิมพ์
2. เพิ่มประสิทธิภาพการทำงานของคอมพิวเตอร์ ด้านการจัดเก็บงานเอกสารที่เป็นกระดาษให้เป็นแฟ้มข้อความมากขึ้น ซึ่งสะดวกและรวดเร็วแก่การค้นหาและแก้ไข

วิจัยและพัฒนาโดย: งานวิจัยเทคโนโลยีภาพ
โครงการวิจัยและพัฒนาโปรแกรมคอมพิวเตอร์แปลงเอกสารภาษาไทย
เก็บเป็นแฟ้มข้อความ ระยะที่ 2
สอบถามข้อมูลทางเทคนิคเพิ่มเติมได้ที่หน้า 249

รหัสผลงาน: BI301-43

ติดต่อสอบถามได้ที่: ฝ่ายพัฒนาธุรกิจและอุตสาหกรรม ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และ
คอมพิวเตอร์แห่งชาติ

โทร. (662) 644-8150..99 ต่อ 610, 656..9

โทรสาร (662) 644-8122, 644-8137..8