234901

การลดมิติถือเป็นงานสำคัญในงานการเรียนรู้ของเครื่องและการรู้จำ เนื่องจากในปัจจุบันนี้เรามีความสามารถในการเก็บข้อมูลทุกๆคุณลักษณะที่เราพอจะนึกถึงได้ ซึ่งนำไปสู่ข้อมูลขนาดใหญ่ที่มีคุณลักษณะมากเกินกว่าที่จะประมวลผลได้ เทคนิคการลดมิติช่วยลดความซ้ำซ้อนระหว่างคุณลักษณะและช่วยขจัดความห่างมากในชุดข้อมูล การเลือกฐานหลักถือเป็นปัญหาเปิดว่าเราจะสามารถลดมิติเพื่อให้ได้เซตย่อยของข้อมูลที่เหมาะที่สุดอย่างอัตโนมัติได้อย่างไร ในงานรู้จำบนชุดข้อมูลรูปภาพนั้นเทคนิคการลดมิติได้รับการพิสูจน์ว่าเป็นตัวประมวลผลก่อนที่สำคัญ เทคนิคแบบสองมิติและสองมิติสองทิศทางเป็นส่วนขยายของเทคนิคการลดมิติสำหรับชุดข้อมูลรูปภาพโดยเฉพาะ เทคนิคแบบสองมิติและสองมิติสองทิศทางให้ค่าความแม่นที่สูงกว่าและมีปริมาณการคำนวณที่น้อยกว่าเมื่อเทียบกับเทคนิคดั้งเดิม เราได้โต้แย้งว่ากระบวนการเลือกฐานหลักในเทคนิคแบบสองมิตินั้นไม่เหมาะที่สุดและไม่ชัดเจนในเทคนิคแบบสองมิติสองทิศทาง เราได้นำเสนอเทคนิคตัวกรองแบบคะแนนเพื่อช่วยทำให้การเลือกฐานหลักเหมาะที่สุดโดยการทำให้เป็นบรรทัดฐานและจัดเรียงฐานหลักใหม่ที่ให้ผลเป็นค่าความแม่นที่สูงขึ้นในการจำแนก นอกจากนี้เรายังนำเสนอวิธีการเลือกฐานหลักใหม่ในเทคนิคแบบสองมิติสองทิศทางสองวิธีได้แก่ วิธีใช้เทคนิคคัดเลือกคุณลักษณะแบบผสมและวิธีอัลกอริทึมเชิงละโมบที่ใช้ควบคู่กับการตัดออก ซึ่งทั้งสองวิธีนี้จะใช้ปริมาณการคำนวณน้อยกว่าวิธีดั้งเดิมที่ใช้การตรวจสอบไขว้ทั้งหมด อีกทั้งยังให้ค่าความแม่นในการจำแนกสูงกว่าวิธีตัดออกโดยอ้างอิงค่าลักษณะเฉพาะที่มักจะติดในหุบเขาของปริภูมิค่าความแม่น เราได้ประยุกต์ใช้แนวคิดกับงานในโลกจริงได้แก่ งานรู้จำใบหน้าและงานรู้จำลายมือ

Dimensional Reduction is an important task in Machine Learning and Pattern Recognition. Nowadays, we have gained an ability to collect every feature in our imaginable aspect which leads to massive data that have too many features to process. Dimensional Reduction techniques reduce the redundancy among features and conquer the sparseness of the data set. Basis Selection is an open problem of how much we should reduce the dimension to be an optimal dataset automatically. In Pattern Recognition tasks of image data sets, techniques for Dimensional Reduction have been proven to be an important preprocessing step. 2D and $(2D)^2$ techniques are the extensions of Dimensional Reduction techniques tailored for image data sets. 2D and $(2D)^2$ techniques provide more classification accuracy with less computational cost compared to original techniques. We argue that the step in Basis Selection of 2D techniques is not optimal and the method of Basis Selection is not obvious in $(2D)^2$ techniques. We propose the score-based filter techniques to further optimize the Basis Selection by performing a normalizing and reordering step which results in higher classification accuracy. Moreover, we also introduce a novel Basis Selection method in $(2D)^2$ techniques based on a greedy algorithm and score metrics which requires less computational cost compared to cross validation based techniques and provides more classification accuracy than eigenvalue-based cut-off techniques which stuck in the valley of accuracy space. We have applied our proposed method to several real world applications, face recognition and hand written recognition.