International Journal of Innovative Computing, Information and Control
Volume **6**, Number **6**, June **2010**

ICIC International ©2010 ISSN 1349-4198
pp. **2645–2656**

HYBRID GLOBAL/LOCAL SEARCH STRATEGIES FOR VQ CODEBOOK GENERATION BASED ON OTSU AND LBG ALGORITHM

Chang-Chin Huang[1], Du-Shiau Tsai[2] and Gwoboa Horng[1]

[1]Department of Computer Science and Engineering
National Chung-Hsing University
Taichung 402, Taiwan
phd9308@cs.nchu.edu.tw; gbhorng@cs.nchu.edu.tw

[2]Department of Information Networking Technology
Hsiuping Institute of Technology
Taichung 412, Taiwan
dstsai@mail.hit.edu.tw

Received January 2009; revised May 2009

Abstract. *Vector quantization algorithms have been extensively used for image compression, pattern recognition, image steganography, image retrieval, and anomaly intrusion detection. For large N_p training vectors and N_c clusters, vector quantization algorithms can hardly find the global optimal classification without requiring a great deal of the squared Euclidean distance calculation. This paper proposes an efficient global division algorithm based on histogram threshold to improve computation time of the squared Euclidean distance from $O(kN_p \log N_c)$ to $O(kN_pN_c)$. The experimental results and comparisons show that the global division algorithm can reduce computational complexity, find better codewords to improve the quality of the codebook and cooperate with the local search algorithm to tune it efficiently.*

Keywords: Vector quantization, Codebook generation, Histogram threshold, Clustering, LBG algorithm, Global optimal, Local optimal

1. **Introduction.** The quantization method is an effective method for high ratio image compression. In the last decade, vector quantization (VQ) algorithms [1-3] have been extensively and successfully used for compression due to their excellent rate-distortion performance and relatively simple structure. Applications in this class are often found in speech and image data compression [1-3], pattern recognition [4,5], image steganography [6,7], image retrieval[8], anomaly intrusion detection[9], and others.

According to the VQ processes, the quality of the reconstructed image highly depends on codewords in the codebook. That is, a good codebook can usually procure high quality images encoded by VQ, whose most important task is the codebook design. The paper aims at advancing a new VQ scheme to satisfy the following guidelines: it should be a scheme with (1) low computational complexity and (2) better reconstructed image quality.

In the pattern recognition literature, the term cluster, not partition, is used, and algorithms to find clusters are called clustering algorithms.

Because the clustering problem in its combinatorial form has been shown to be NP-complete [1], feasible suboptimal solutions by heuristic or approximate generation algorithms are needed. Generally, several approaches, such as k-mean, hierarchical, and optimization, have been used for designing clustering algorithms.

Before introducing the clustering algorithms, notations in Table 1 are used throughout the paper.