

# A Parallel Processing Framework using MapReduce for Content-Based Image Retrieval

Anucha Tungkasthan<sup>1</sup>, Wichian Premchaiswadi<sup>2</sup>  
Graduate School of Information Technology  
Siam University  
Bangkok, Thailand  
aimdala@hotmail.com<sup>1</sup>, wichian@siam.edu<sup>2</sup>

**Abstract**—In this paper, a Hadoop MapReduce framework is presented in order to perform distributed processing used for CBIR system. Moreover, Hadoop MapReduce framework is used with the intention of increasing the performance of two main functionalities of data insertion and query processing. Therefore, the main objective of the study is distribution of the image data over a large number of nodes. Some of the techniques used in the paper includes: image indexing and retrieval, parallel processing of indexing, and comparing the similarity of retrieved images

**Keywords**—Image Retrieval, CBIR Systems, MapReduce, Distributed Processing, Image Processing