Constant time median filtering of extra large images using Hadoop

Mohammadreza Azodinia[a], Vahid Farrokhi[b], Andras Hajdu[c]

[a]University of Debrecen
m.r.azodinia@inf.unideb.hu
[b]University of Debrecen
vahid.farrokhi@inf.unideb.hu
[c]University of Debrecen
hajdu.andras@inf.unideb.hu

Abstract

The spatial resolution of remote sensing images such as MRI, CT and PET is constantly increasing and analyzing these images in real time is a challenging task. But this limits the efficiency of many image processing algorithms. Among different efficient image processing algorithms, median filtering is a principal element in many image processing situations which manages to reduce the noise while preserving the edges. Median Filtering in Constant Time (MFCT) is a simple yet fastest median filtering algorithm which can handle N-dimensional data in fields like medical imaging and astronomy. With trend toward the median filtering of large images and proportionally large kernels, Hadoop MapReduce (a popular big data processing engine) can be applied and utilized. MapReduce provides the simplicity of defining the map and reduce functions while the framework takes care of parallelization and failover automatically.

Hence, in this paper we discuss on possibility of the incorporation of MFCT algorithm with Hadoop MapReduce framework to improve the performance of processing of extra large images.

Keywords: median filtering, MFCT, MapReduce, Hadoop, parallelization

MSC: 68-06, 68U10