

Utilização do PySpark Para Detecção de Bordas e Borrramento em Imagens

Projeto 2 - Grupo 05



Grupo 5 - Integrantes


Lucas Alves Racoci 156331

Luiz Fernando Rodrigues da Fonseca 156475

Rafael Zane 157079

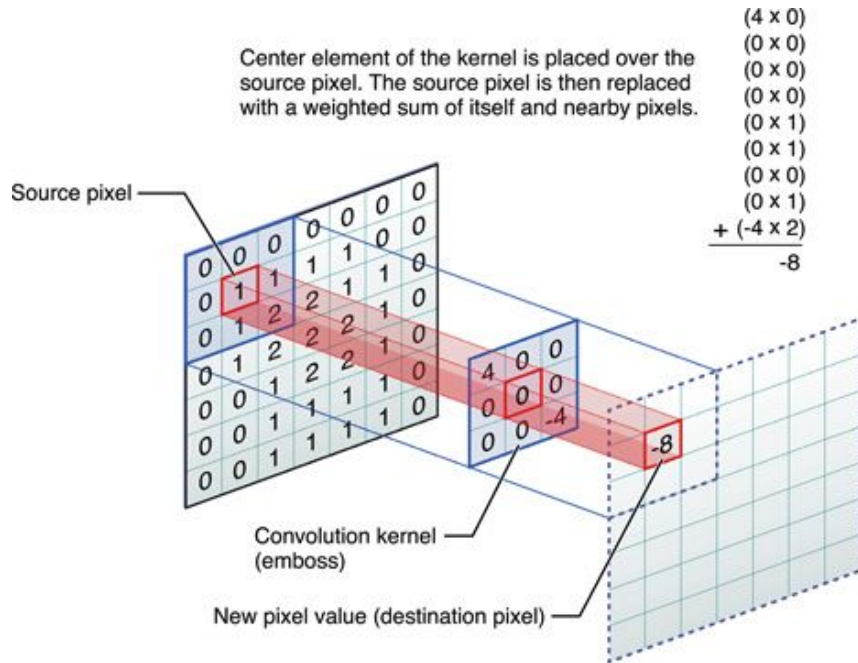
Rodrigo Noronha Máximo 157209

Introdução

- 
- Convolução
 - Explicação básica do funcionamento do Sobel e da Mediana
 - Esquema da aplicação implementada
 - Conclusão - Comparação entre o Spark e o Hadoop HIPI



Convolução




| | | | | |
|-----------------|-----------------|-----------------|---|---|
| 1 _{x1} | 1 _{x0} | 1 _{x1} | 0 | 0 |
| 0 _{x0} | 1 _{x1} | 1 _{x0} | 1 | 0 |
| 0 _{x1} | 0 _{x0} | 1 _{x1} | 1 | 1 |
| 0 | 0 | 1 | 1 | 0 |
| 0 | 1 | 1 | 0 | 0 |

Image

| | | |
|---|--|--|
| 4 | | |
| | | |
| | | |


Convolved
Feature

Sem Sobreposição de Blocos



| | | | | | | | |
|----|----|----|----|----|----|----|----|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |

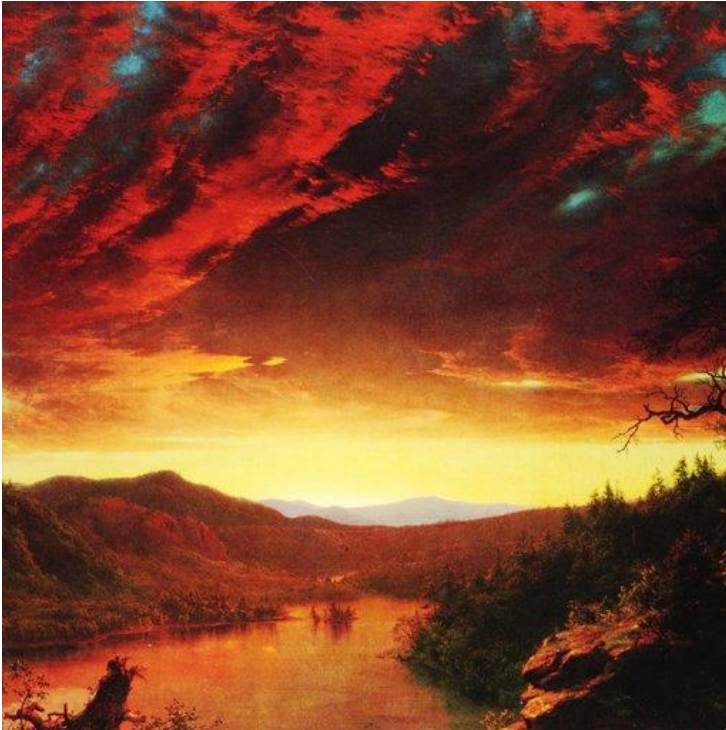
Com Sobreposição de Blocos



| | | | | | | | |
|----|----|----|----|----|----|----|----|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
| | | | | | | | |

Filtro da Mediana

- Efeito de blur



Filtro



Filtro Sobel

- Detecção de bordas



Filtro



Filtro Sobel

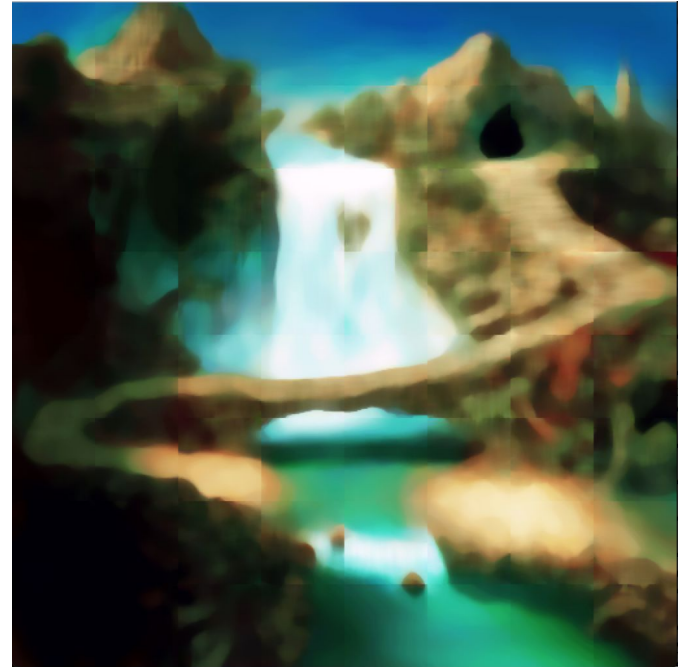
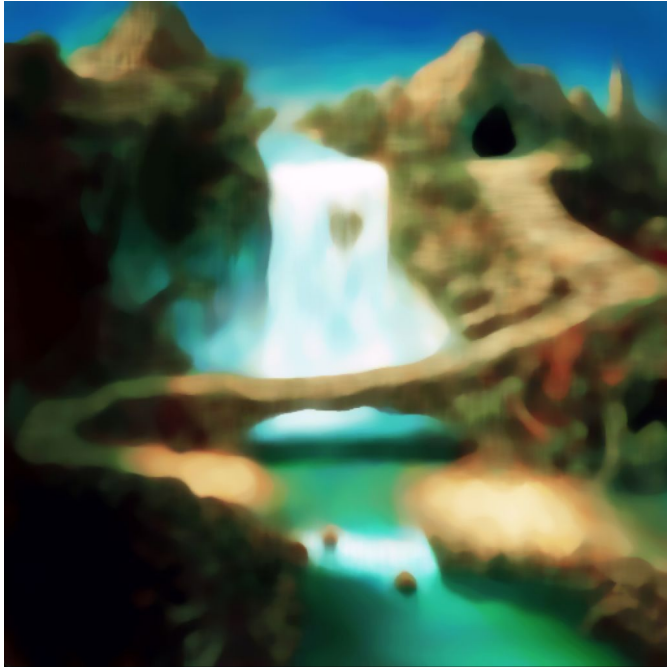
- Detecção de bordas de uma imagem A

$$\mathbf{G}_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * \mathbf{A} \quad \text{e} \quad \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$

$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2}$$

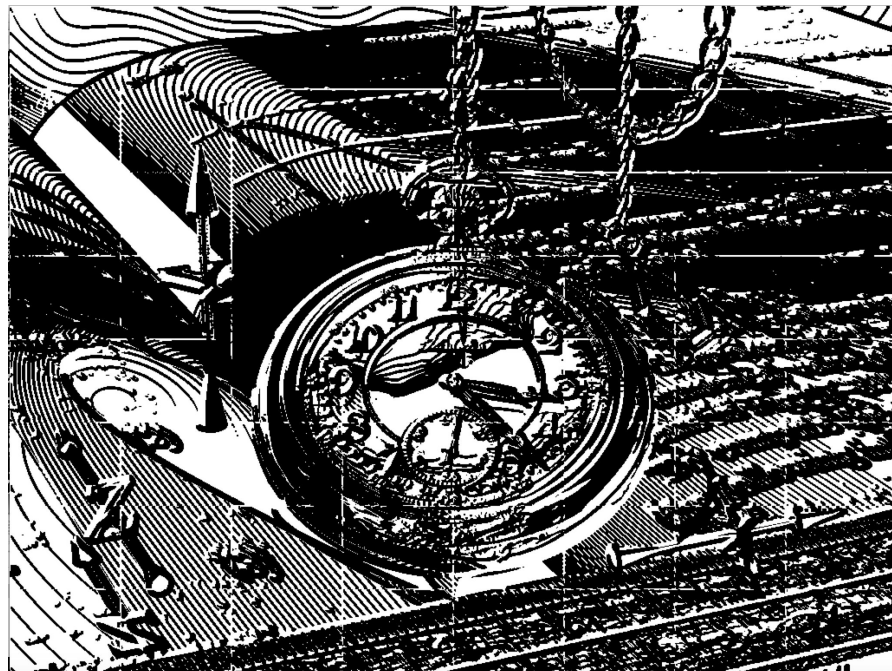
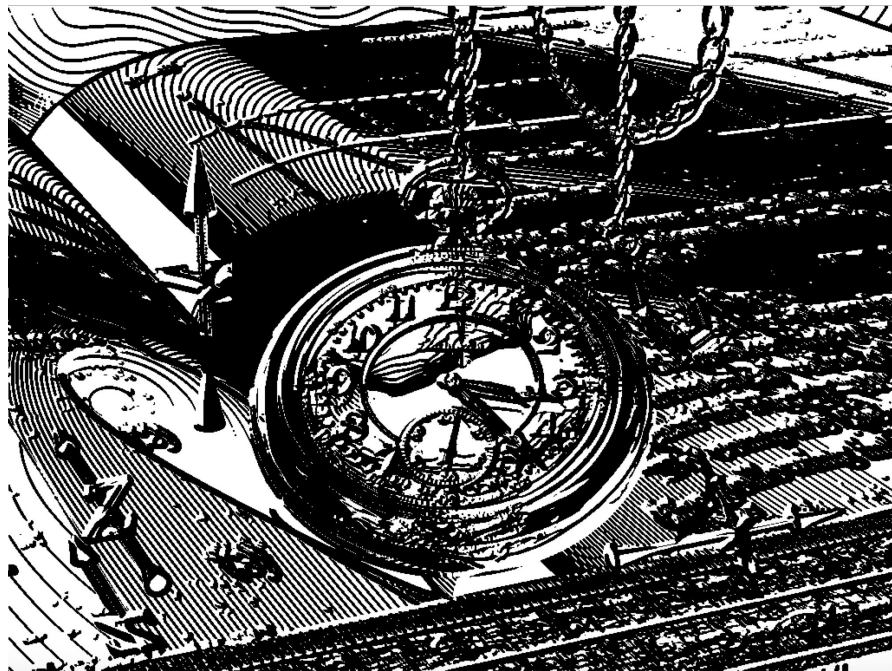
Problemas da Sobreposição na Mediana

- Efeitos indesejados

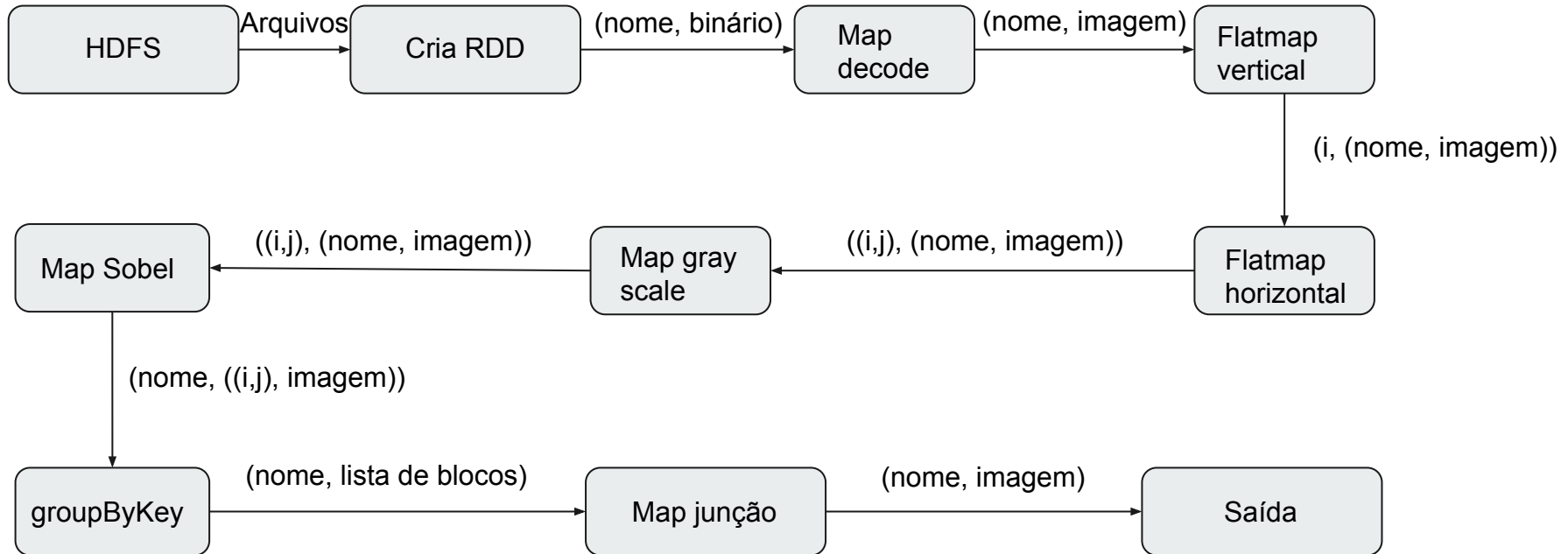


Problemas da Sobreposição no Sobel

- Efeitos indesejados



Esquema de Processamento



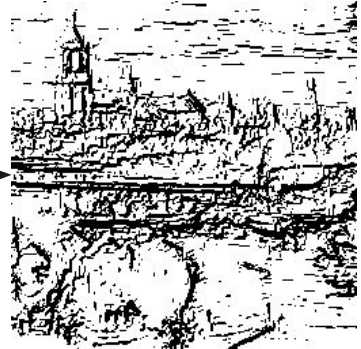
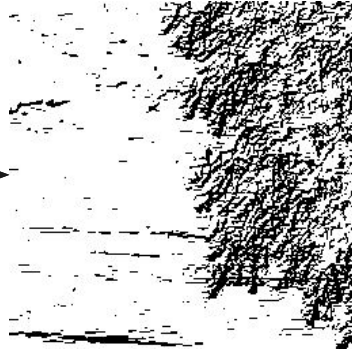
Esquema de Processamento



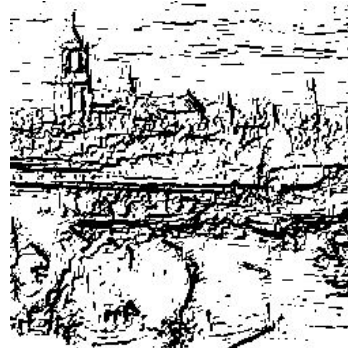
Esquema de Processamento



Resultados



Resultados



Conclusão



- Comparação Hadoop HIPI - Spark
- O PySpark roda com várias bibliotecas que o Hadoop HIPI não possui acesso, ou que possui um acesso incompleto (OpenCV)
- RDDs diversificam muito as possibilidades de processamento paralelo, fornecendo operações genéricas que facilitam o tratamento de dados, além de fazer as operações na memória
- O PySpark também possui um modo iterativo que é melhor para o desenvolvimento
- A integração PySpark com o HDFS é feita de maneira mais fácil, mas ainda algumas funções estão em desenvolvimento