

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO



PROCESSAMENTO DIGITAL DE IMAGENS - RELATÓRIO DE  
TRABALHO SOBRE DETECÇÃO DE MUDANÇAS

---

Flávio Roberto dos Santos Coutinho (2019679021)

30 de Junho de 2019

# 1 Introdução

O trabalho de detecção de mudanças no *dataset* ONERA<sup>1</sup>, obtido por meio do satélite Sentinel-2, foi realizado usando duas técnicas e, então, as acurácias obtidas com cada uma foram comparadas<sup>2</sup>.

A primeira técnica usada foi *image differencing* [1], uma técnica algébrica simples que prevê a subtração dos tempos  $t_1$  e  $t_2$  e a classificação de cada pixel como mudança ou não-mudança de acordo com a diferença ser maior ou não que um *threshold* definido.

Complementarmente, foi implementada outra solução, baseada na redução de dimensionalidade das imagens em  $t_1$  e em  $t_2$  usando PCA [1], seguido do cálculo de descritores de textura com LBP [2], da segmentação em superpixels com SLIC [3] e da classificação de cada superpixel em mudança ou não-mudança de acordo com a distância euclidiana das características de cada segmento em  $t_1$  e  $t_2$ .

## 2 Desenvolvimento e Metodologia

A solução que usa *image differencing* foi implementada conforme metodologia de Rivera (2005) [4] (Figura 1a) e a segunda solução foi baseada em PCA, LBP e segmentação com SLIC (Figura 1b).

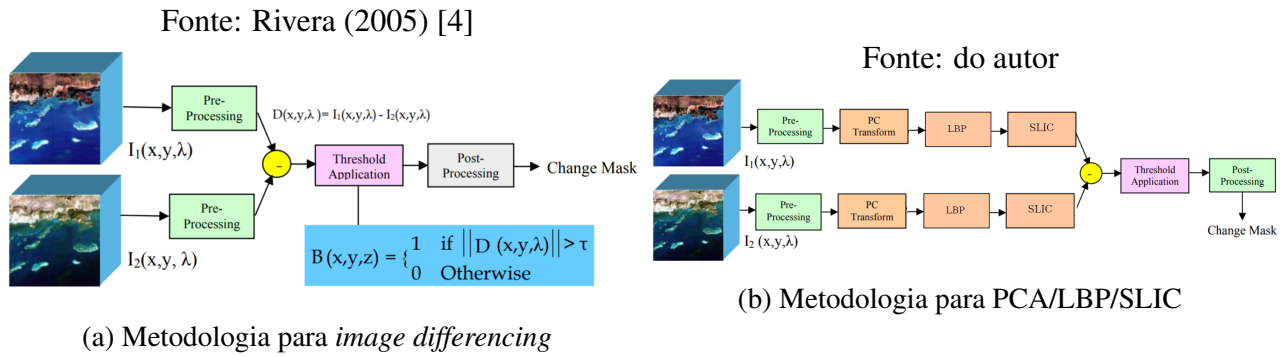


Figura 1: Esquema da metodologia para *image differencing* e para PCA/LBP/SLIC

Em ambos os casos, as imagens foram pre-processadas para que todas as bandas tivessem as dimensões da maior banda. Esse ajuste de *upscaling* foi feito usando interpolação bicúbica. Assumiu-se que as imagens de  $t_1$  e  $t_2$  já estavam corretamente registradas.

Sobre a **primeira solução**, foi usado um valor de *threshold* de 3,5% da intensidade de cada pixel - caso a diferença entre  $t_1$  e  $t_2$  seja maior que esse valor, esse pixel é considerado como mudança. Esse valor foi determinado experimentalmente.

A **segunda solução** seguiu as etapas de *i*) redução da dimensionalidade (13 bandas) de cada  $t$  usando PCA e gerando uma nova imagem com os 3 componentes principais, *ii*) cálculo do descritor de textura LBP da imagem gerada via PCA, *iii*) segmentação da imagem  $t_1$  usando SLIC *iv*) comparação das características de cada segmento em  $t_1$  e  $t_2$  usando distância euclidiana do histograma da imagem LBP e *v*) classificação em mudança ou não-mudança de cada segmento de acordo com um *threshold* específico para cada cidade.

A (i) redução de dimensionalidade foi feita para acelerar o processamento ao descartar informação redundante, de baixa variância, nas 13 bandas de cada pixel. As imagens foram reduzidas de 13 bandas para 3, mantendo em média 89% da variância total presente nas bandas originais. Foi gerada uma imagem PCA para  $t_1$  e outra para  $t_2$ .

As imagens PCA para  $t_1$  e  $t_2$  tiveram (ii) descritores de textura LBP <sup>enunc.2</sup> <sup>3</sup> calculados usando o método *nri\_uniform* [2] que é invariante ao valor de cinza. Esse método foi usado em detrimento

<sup>1</sup>Enunciado: [https://github.com/fegemo/change-detection/blob/master/EnunciadoTP2\\_change.pdf](https://github.com/fegemo/change-detection/blob/master/EnunciadoTP2_change.pdf)

<sup>2</sup>Código, resultados e relatório: <https://github.com/fegemo/change-detection>

<sup>3</sup>Referente ao item 2 do enunciado do trabalho

de outros que são invariantes à rotação porque assumiu-se que as imagens  $t_1$  e  $t_2$  já estavam registradas. Os descritores foram gerados separadamente por banda. LBP foi usado em vez de GLCM porque a matriz de coocorrência precisaria ser gerada para cada segmento (criado no passo seguinte), e a imagem de padrões binários poderia ser gerada apenas uma vez.

A imagem  $t_1$  foi (iii) segmentada em superpixels com SLIC [enunc.1](#) usando *compactness* 20 e um máximo de 2.000 segmentos. Os mesmos segmentos (de  $t_1$ ) foram usados para comparar  $t_1$  e  $t_2$ . Foi gerado um descritor de textura de cada segmento baseado no (iv) [enunc.2](#) histograma da imagem LBP dos pixels que compunham cada segmento e a distância euclidiana entre esse descritor de  $t_1$  e  $t_2$  foi calculada. Em seguida, os segmentos cuja distância foi maior do que um valor de *threshold* foram (v) classificados como mudança [enunc.3](#). Esse *threshold* foi definido por cidade e foi calculado como a média das distâncias dos segmentos mais dois desvios padrões. Esse valor foi obtido experimentalmente.

Depois disso, foi gerado um [enunc.4](#) mapa binário de mudanças mostrando o resultado de todos os segmentos e a acurácia do mapa quanto ao *ground truth* foi avaliada [enunc.5](#).

### 3 Experimentos e Testes

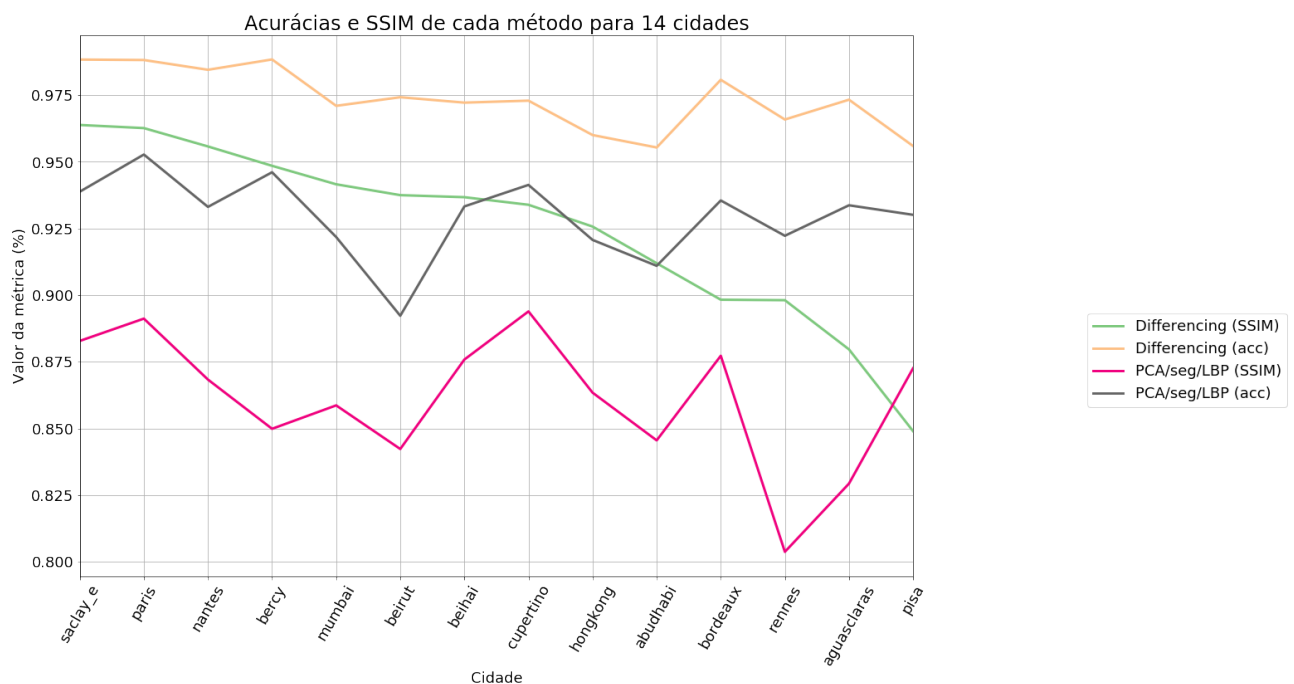
Ambas soluções geram mapas binários que indicam, para cada pixel, se houve mudança de  $t_1$  para  $t_2$  ou não. A comparação dos mapas com a *ground truth* foi feita por meio das métricas *structural similarity index* (SSIM) e acurácia, sendo esta definida pelas Equações 1 e 2, onde  $a$  é a quantidade de pixels de uma imagem e  $d$  é a quantidade de pixels diferentes entre a  $img_1$  e  $img_2$ .

$$acc = (a - d)/a \quad (1)$$

$$d = \|abs(img_1 - img_2) > 0\| \quad (2)$$

Os valores de acurácia e SSIM obtidos para as 14 cidades cujo *ground truth* era conhecido podem ser vistos numericamente na Tabela 1 ou visualizados no gráfico da Figura 2. Os mapas binários com as diferenças detectadas pelas duas técnicas para as 14 cidades podem ser vistos na Figura 3.

Figura 2: Gráfico com SSIM e acurácia de cada técnica para as 14 cidades com *ground truth*



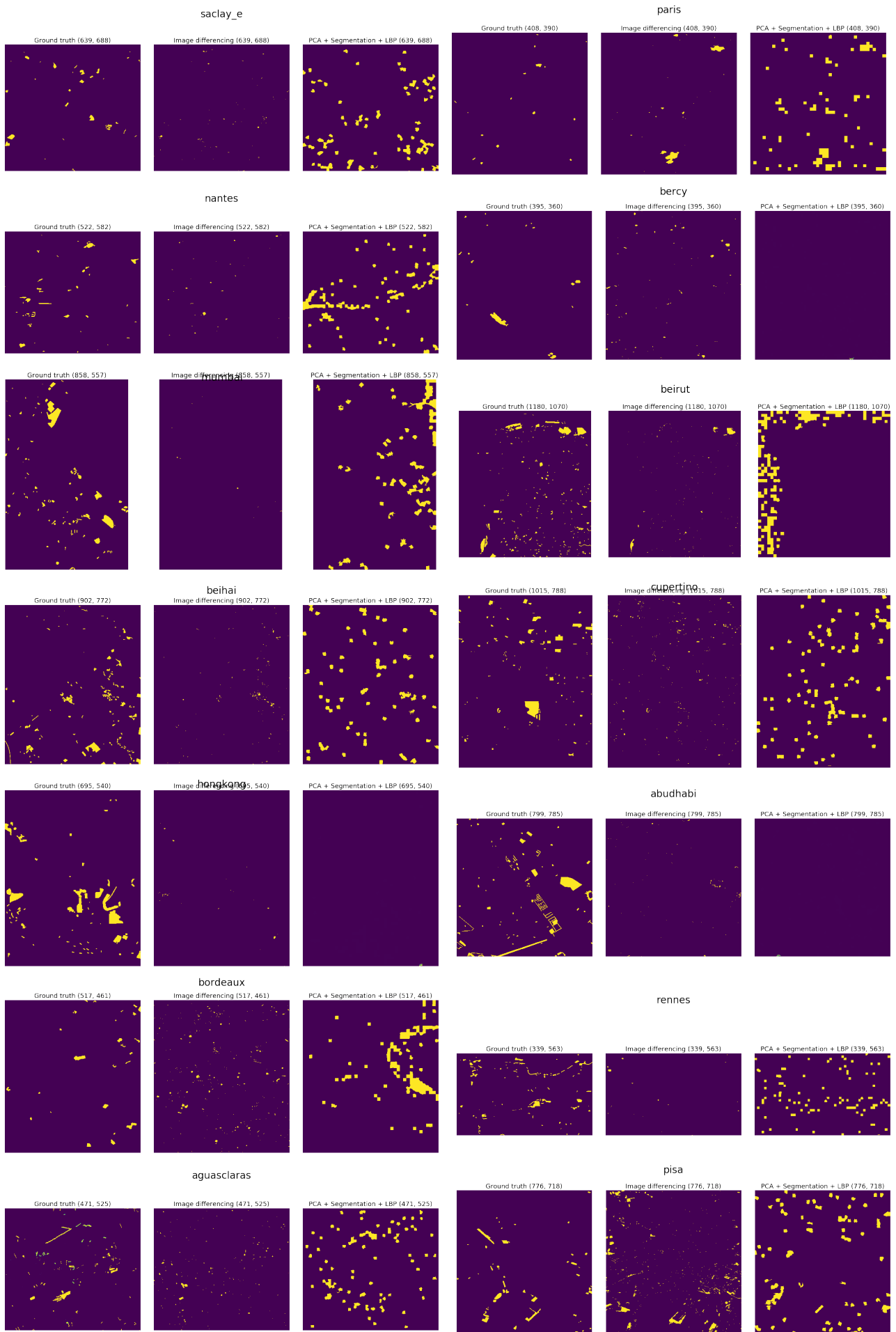


Figura 3: Mapas binários e *ground truth* das 14 cidades

Tabela 1: SSIM e Acurácia das 14 cidades com *ground truth*

Cidade	Image Differencing		PCA/seg/LBP	
	SSIM	Acurácia	SSIM	Acurácia
saclay_e	0,9638	0,9883	0,8828	0,9388
paris	0,9626	0,9881	0,8912	0,9527
nantes	0,9557	0,9845	0,8683	0,9331
bercy	0,9485	0,9883	0,8499	0,9460
mumbai	0,9415	0,9710	0,8586	0,9218
beirut	0,9375	0,9742	0,8423	0,8922
beihai	0,9367	0,9721	0,8758	0,9332
cupertino	0,9339	0,9729	0,8939	0,9413
hongkong	0,9257	0,9600	0,8634	0,9207
abudhabi	0,9119	0,9553	0,8455	0,9110
bordeaux	0,8983	0,9807	0,8772	0,9355
rennes	0,8981	0,9658	0,8037	0,9222
aguasclaras	0,8796	0,9733	0,8293	0,9337
pisa	0,8490	0,9559	0,8726	0,9301
<b>Média</b>	0,9245	0,9736	0,8611	0,9295
<b>Desvio</b>	0,0334	0,0115	0,0250	0,0153

## 4 Conclusão

Os resultados usando a solução *image differencing* foram melhores do que usando PCA/LBP/SLIC. Motivos para o pior resultado da segunda técnica pode ter sido a definição de parâmetros experimentais. Outro potencial problema é a detecção ser feita no segundo caso considerando um superpixel inteiro como mudança ou não-mudança.

## Referências

- [1] D. Lu, P. Mausel, E. Brondízio, and E. Moran. Change detection techniques. *International Journal of Remote Sensing*, 25(12):2365–2401, 2004.
- [2] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (12):2037–2041, 2006.
- [3] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11):2274–2282, 2012.
- [4] Vanessa Ortiz Rivera. *Hyperspectral change detection using temporal Principal component analysis*. PhD thesis, Citeseer, 2005.