

Tarefa II:

Reconstrução de textura guiada por *patch* de alta resolução.

André Pigatto, Israel de Oliveira, Rafael Licursi
PPGEE (Programa de Pós-graduação em Engenharia Elétrica)

Universidade Federal do Rio Grande do Sul

Visão Computacional (ELE311)
Prof. Jacob Scharcanski

13 de julho, 2017, Porto Alegre

M. El Gheche, J. F. Aujol, Y. Berthoumieu and C. A. Deledalle, "Texture Reconstruction Guided by a High-Resolution Patch," in IEEE Transactions on Image Processing, vol. 26, no. 2, pp. 549-560, Feb. 2017.

Contents

1 Introdução

2 Método proposto

- A: Wasserstein Distance to the High-Resolution Patch
- B: Nonlocal Regularization
- Algoritmo
- Inconsistências

3 Resultados

- Métricas
- Imagens
- Métricas

4 Conclusões

- O artigo escolhido [Gheche et al. 2017] apresenta uma proposta de reconstrução de texturas guiada por uma amostra (*patch*) de alta resolução.
- É proposto um algoritmo baseado em um problema de otimização usando um método primal-dual.
- Formulação de um custo envolvendo dois termos: as distâncias de Wasserstein do histograma da amostra de alta resolução, do gradiente e do laplaciano dessa amostra; uma regularização NLTV usando a mesma amostra.
- Os métodos anteriores usam a regularização NLTV sem levar em conta um amostra de alta resolução e uma forma indireta das distâncias de Wasserstein com treinamento prévio.

Método proposto I

- Imagem original $X \in \mathbb{R}^{N_1 \times N_2}$, vetorizada como $\mathbf{x} = \text{vec}(X) \in \mathbb{R}^N$, $N = N_1 N_2$:

$$z_1 = DB\mathbf{x} + \eta \quad (1)$$

$$z_2 = M\mathbf{x} \quad (2)$$

sendo $D \in \mathbb{R}^{Q \times N}$ matriz de amostragem, $B \in \mathbb{R}^{Q \times N}$ responsável por aplicar um efeito de borramento, $\eta \in \mathbb{R}^Q$ ruído aditivo gaussiano com média nula e variância τ .

- Deseja-se obter $x^* = \{x \mid \min \|x - \mathbf{x}\|\}$ a partir de z_1 e z_2 :

$$\{x^*\} = \arg \min_{x \in \mathbb{R}^N} \|DB\mathbf{x} - z_1\|^2 + H(x, z_2) + \lambda R(x, z_2) \quad (3)$$

$$s.a. \quad Mx = z_2 \quad (4)$$

A: Wasserstein Distance to the High-Resolution Patch

- Distância de Wasserstein entre dois histogramas normalizados ν_u e ν_ν , das imagens u e ν , com Σ_N todos as possíveis permutações de vetores \mathbb{R}^N :

$$\mathcal{W}_2^2(\nu_u, \nu_\nu) = \min_{\sigma \in \Sigma_N} \|u - \nu \circ \sigma\|^2$$

no caso de imagens em escala de cinza: $\sigma^* = \sigma_\nu \circ \sigma_u^{-1}$.

- Gradiente de $\mathcal{W}_2^2(\cdot)$:

$$\nabla_2 \mathcal{W}_2^2(\nu_u, \nu_\nu) = 2(u - \nu \circ \sigma_\nu \circ \sigma_u^{-1})$$

A: Wasserstein Distance to the High-Resolution Patch



u



v



$v \circ \sigma_v$



$v \circ \sigma_v \circ \sigma_u^{-1}$

A: Wasserstein Distance to the High-Resolution Patch

$$H(x, z_2) = \sum_{s=1}^4 a_s \mathcal{W}_2^2(\nu_{x_s}, \nu_{z_{2s}})$$

com $x_s = L_s x$ e $z_{2s} = L_s z_2$.

- L_1 : matriz identidade ($N_1 = N$).
- L_2 : concatenação da diferença horizontal e vertical ($N_2 = 2N$).
- L_3 : concatenação da diferença diagonal ($N_3 = 2N$).
- L_4 : operador laplaciano isotrópico.
- Os operadores L_3 e L_4 representam o operador gradiente.

Como obter as matrizes L_s ? Os autores não citam referências.

B: Nonlocal Regularization

NLTV proposto:

$$R(x, z_2) = \sum_{i=1}^N \sqrt{\sum_{j \in \mathcal{N}_i} w_{ij} (x_i - z_{2j})^2}$$

sendo \mathcal{N}_i um conjunto de suporte referente à z_2 .

$$w_{ij} = e^{-\delta \|p_{\bar{j}i}(z_2) - p_j(z_2)\|^2}$$

com $p_j(\cdot)$ um block centrado no pixel j e δ uma constante positiva (nenhum valor é apresentado no artigo).

$$\bar{j}^i = \arg \min_{m \in 1, \dots, M} \|p_i(\hat{z}_1) - p_m(z_2)\|^2$$

com \hat{z}_1 sendo uma estimativa de \mathbf{x} .

Algorithm 1 FBPD [54]

INITIALIZATION

$$\left[\begin{array}{l} \text{Choose } (x^{[0]}, y^{[0]}) \in \mathbb{R}^n \times \mathbb{R}^{Kn} \\ \text{set } \tau > 0 \text{ and } \omega > 0 \text{ such that} \\ \tau (\beta/2 + \omega \lambda \|\mathbf{T}\|^2) < 1 \end{array} \right.$$

FOR $l = 0, 1, \dots$

$$\left[\begin{array}{l} \hat{x}^{[l]} = \nabla f(x^{[l]}) + \mathbf{T}^\top y^{[l]} \\ x^{[l+1]} = P_{\{\mathbf{M} \cdot = \mathbf{z}^{(2)}\}}(x^{[l]} - \tau \hat{x}^{[l]}) \\ \hat{y}^{[l]} = \mathbf{T}(2x^{[l+1]} - x^{[l]}) \\ y^{[l+1]} = \text{prox}_{\omega \lambda \|\cdot\|_{1,2}^*}(y^{[l]} + \omega \hat{y}^{[l]}) \end{array} \right.$$

- The nonlocal regularization can be expressed as the $\ell_{1,2}$ -norm composed with a discrete difference operator, yielding

$$R(x, z^{(2)}) = \|\mathbf{T}x\|_{1,2}, \quad (15)$$

where

$$\mathbf{T}x = \begin{bmatrix} [w_{1,j}(x_1 - z_j^{(2)})]_{j \in \mathcal{N}_1} \\ \vdots \\ [w_{N,j}(x_N - z_j^{(2)})]_{j \in \mathcal{N}_N} \end{bmatrix} \begin{matrix} \} \in \mathbb{R}^K \\ \vdots \\ \} \in \mathbb{R}^K. \end{matrix} \quad (16)$$

The proximity operator of the $\ell_{1,2}$ -norm, that is $\text{prox}_{\|\cdot\|_{1,2}}$ in Algorithm 1, can be found in [62].

Algoritmo: $P_{\{M \cdot = z_2\}}(\cdot)$

- A $P_{\{M \cdot = z_2\}}(\cdot)$ associa a restrição do problema de otimização:

$$P_{\{M \cdot = z_2\}}(x) = x + M^T(z_2 - Mx)$$

- $z = Mx$ extrai uma parte (*patch*) de x , e $\tilde{x} = M^T z$ retorna \tilde{x} nulo exceto nos respectivos pixels de z .

Algoritmo: $\nabla f(x)$

- $\nabla_x (\|DB\mathbf{x} - z_1\|^2 + H(x, z_2))$

$$\nabla f(x) = 2D^T B^T (DBx - z_1) + 2 \sum_{s=1}^4 a_s L_s^T (L_s x - z_{2s} \circ \sigma_{z_{2s}} \circ \sigma_{L_s x}^{-1})$$

Obs.: $\nabla f(x)$ é β -Lipschitz, entretanto não garante que o algoritmo 1 seja estável. Conforme experienciado pelos autores (e por nós) sempre converge. Ou, como escrito no artigo: "(...) nós observamos que sempre converge para uma solução estável".

- A expressão $z_{2s} \circ \sigma_{z_{2s}} \circ \sigma_{L_s x}^{-1}$ é obtida transferindo o histograma de z_{2s} para $L_s x$.

Algoritmo: $\text{prox}_{\omega\lambda\|\cdot\|_{1,2}^*}(y)$

- Operador de proximidade com função $\phi(\cdot)$ (*ϕ -function proximity operator*):

$$\text{prox}_{\phi}(y) = \arg \min_{z \in \mathbb{R}^N} \phi(z) + \frac{1}{2} \|z - y\|^2 \quad (\forall y \in \mathbb{R}^N)$$

- Função usada:

$$\text{prox}_{\omega\lambda\|\cdot\|_{1,2}^*}(y) = \arg \min_{z \in \mathbb{R}^N} \omega\lambda\|z\|_{1,2} + \frac{1}{2} \|z - y\|^2$$

Algorithm 1 FBPD [54]

INITIALIZATION

$$\left[\begin{array}{l} \text{Choose } (x^{[0]}, y^{[0]}) \in \mathbb{R}^n \times \mathbb{R}^{Kn} \\ \text{set } \tau > 0 \text{ and } \omega > 0 \text{ such that} \\ \tau (\beta/2 + \omega \lambda \|\mathbf{T}\|^2) < 1 \end{array} \right.$$

FOR $l = 0, 1, \dots$

$$\left[\begin{array}{l} \hat{x}^{[l]} = \nabla f(x^{[l]}) + \mathbf{T}^\top y^{[l]} \\ x^{[l+1]} = P_{\{\mathbf{M} \cdot = \mathbf{z}^{(2)}\}}(x^{[l]} - \tau \hat{x}^{[l]}) \\ \hat{y}^{[l]} = \mathbf{T}(2x^{[l+1]} - x^{[l]}) \\ y^{[l+1]} = \text{prox}_{\omega \lambda \|\cdot\|_{1,2}^*}(y^{[l]} + \omega \hat{y}^{[l]}) \end{array} \right.$$

- The nonlocal regularization can be expressed as the $\ell_{1,2}$ -norm composed with a discrete difference operator, yielding

$$R(x, z^{(2)}) = \|\mathbf{T}x\|_{1,2}, \quad (15)$$

where

$$\mathbf{T}x = \begin{bmatrix} [w_{1,j}(x_1 - z_j^{(2)})]_{j \in \mathcal{N}_1} \\ \vdots \\ [w_{N,j}(x_N - z_j^{(2)})]_{j \in \mathcal{N}_N} \end{bmatrix} \begin{matrix} \} \in \mathbb{R}^K \\ \vdots \\ \} \in \mathbb{R}^K. \end{matrix} \quad (16)$$

The proximity operator of the $\ell_{1,2}$ -norm, that is $\text{prox}_{\|\cdot\|_{1,2}}$ in Algorithm 1, can be found in [62].

- Norma $L_{p,q}$:

$$\|A\|_{p,q} \left(\sum_{i=1}^N \left(\sum_{j=1}^K |a_{ij}|^q \right)^{\frac{p}{q}} \right)^{\frac{1}{p}}$$

- Norma $L_{1,2}$:

$$\|A\|_{1,2} \left(\sum_{i=1}^N \left(\sum_{j=1}^K a_{ij}^2 \right)^{\frac{1}{2}} \right)^{\frac{1}{1}}$$

- $R(\cdot)$ como uma norma $L_{1,2}$:

$$R(x, z_2) = \sum_{i=1}^N \sqrt{\sum_j^K w_{ij} (x_i - z_{2j})^2} \quad (5)$$

$$= \sum_{i=1}^N \sqrt{\sum_j^K (\sqrt{w_{ij}} (x_i - z_{2j}))^2} \quad (6)$$

$$= \|\mathbf{T}x\|_{1,2} \quad (7)$$

$$(\mathbf{T}x)_{ij} = (\mathbf{T}(x))_{ij} = \sqrt{w_{ij}} (x_i - z_{2j}), \quad i = 1, \dots, N \text{ e } j = 1, \dots, K$$

Portanto, $\mathbf{T}x : \mathbb{R}^N \mapsto \mathbb{R}^{N \times K}$.

Inconsistências:

- 1 Como $\mathbf{T}^T y \in \mathbb{R}^N$, considerando $\mathbf{T}x : \mathbb{R}^N \mapsto \mathbb{R}^{N \times K}$?
- 2 Como $\mathbf{T}^T y \in \mathbb{R}^N$, considerando $\mathbf{T}x : \mathbb{R}^N \mapsto \mathbb{R}^{KN}$?
- 3 Se $y \in \mathbb{R}^{KN}$ e $x \in \mathbb{R}^N$, poderia $\mathbf{T}(y)$?
- 4 $y = \mathbf{T}x$, apenas se considerar $\mathbf{T}x : \mathbb{R}^N \mapsto \mathbb{R}^{KN}$.

Implementado sem sucesso:

- $\mathbf{T}x : \mathbb{R}^N \mapsto \mathbb{R}^{KN}$, quando $x \in \mathbb{R}^N$.
- $\mathbf{T}y : \mathbb{R}^{KN} \mapsto \mathbb{R}^N$, quando $y \in \mathbb{R}^{KN}$.

$$(\mathbf{T}x)_{i \times j} = w_{ij} (x_{i \times j} - z_{2j})^2, \quad i = 1, \dots, N \text{ e } j = 1, \dots, K$$

Implementado com sucesso:

- $\mathbf{T}x : \mathbb{R}^N \mapsto \mathbb{R}^{K \times N}$, $y \in \mathbb{R}^K$.
- $\nabla_x \|\mathbf{T}x\|_{1,2} = (\mathbf{T}x)y$

$$(\mathbf{T}x)_{ij} = w_{ij} (x_i - z_{2j})^2, \quad i = 1, \dots, N \text{ e } j = 1, \dots, K$$

Algorithm 1 FBPD [24]

INITIALIZATION

Choose $(x^{[0]}, y^{[0]}) \in \mathbb{R}^n \times \mathbb{R}^{Kn}$
set $\tau > 0$ and $\omega > 0$ such that $\tau(\beta/2 + \omega) < 1$

FOR $l = 0, 1, \dots$

$\hat{x}^{[l]} = \nabla f(x^{[l]}) + y^{[l]}$
 $x^{[l+1]} = P_{\{M \cdot = z^{(2)}\}}(x^{[l]} - \tau \hat{x}^{[l]})$
 $\hat{y}^{[l]} = (2x^{[l+1]} - x^{[l]})$
 $y^{[l+1]} = y^{[l]} + \omega \hat{y}^{[l]} - P_{C_2}(y^{[l]} + \omega \hat{y}^{[l]})$

Problema *primal-dual* [Condat 2013]

$$\text{Find } \hat{x} \in \arg \min_{x \in \mathcal{X}} [F(x) + G(x) + H(Lx)]$$

Com $x \in \mathbb{R}^N$ e $y \in \mathbb{R}^K$, $L : \mathbb{R}^N \mapsto \mathbb{R}^K$.

$$\text{Find } \hat{y} \in \arg \min_{y \in \mathcal{Y}} [(F + G)^*(-L^*y) + H^*(y)]$$

$$\tilde{x}_{n+1} := \text{prox}_{\tau G}(x_n - \tau(\nabla F(x_n) + e_{F,n}) - \tau L^* y_n) + e_{G,n},$$

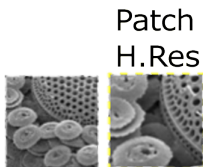
$$\tilde{y}_{n+1} := \text{prox}_{\sigma H^*}(y_n + \sigma L(2\tilde{x}_{n+1} - x_n)) + e_{H,n},$$

$$(x_{n+1}, y_{n+1}) := \rho_n(\tilde{x}_{n+1}, \tilde{y}_{n+1}) + (1 - \rho_n)(x_n, y_n),$$

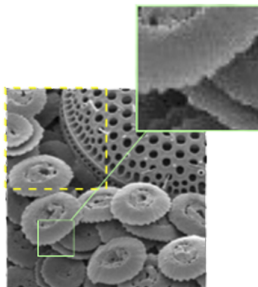
Sendo $e_{\{\cdot\},\{\cdot\}}$ as respectivas inexatidões (detrativos) computacionais.

- SSIM: índice de similaridade estrutural (*structural similarity index*)
- PSNR: relação sinal-ruído de pico (*peak signal-to-noise ratio*)
- MSE: erro quadrático médio (*Mean Squared Error*)
- Erro estatístico proposto pelos autores:

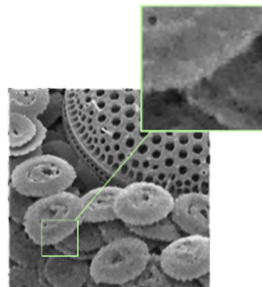
$$\text{Err}_{\bar{x}} = \sum_{s=1}^4 \frac{a_s}{N_s} \mathcal{W}_2^2(\nu_x, \nu_{\bar{x}})$$



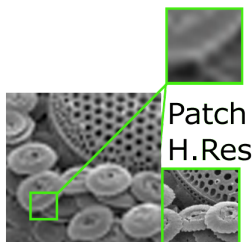
(a) z1: Artigo
Fig. Low Res.



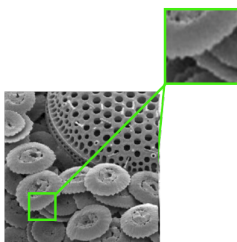
(b) Artigo
Fig. Original.



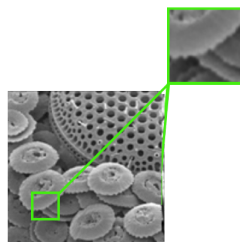
(c) x: Artigo
Fig. Reconstituída



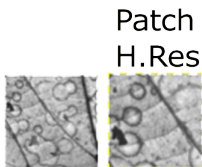
(d) z1: Nosso
Fig. Low Res.



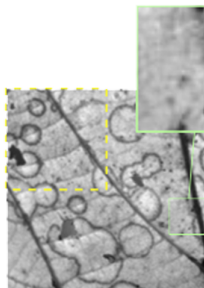
(e) Nosso
Fig. Original.



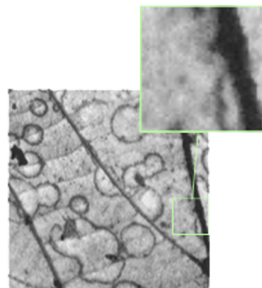
(f) x: Nosso
Fig. Reconstituída



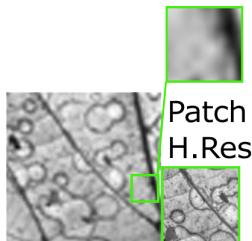
(a) z1: Artigo
Fig. Low Res.



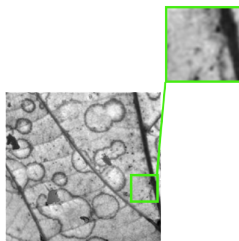
(b) Artigo
Fig. Original.



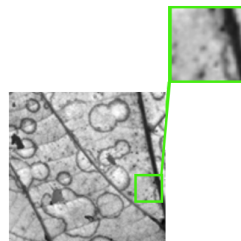
(c) x: Artigo
Fig. Reconstituída



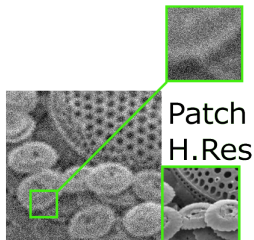
(d) z1:Nosso
Fig. Low Res.



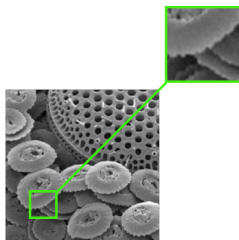
(e) Nosso
Fig. Original.



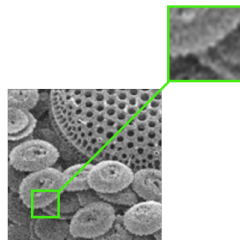
(f) x: Nosso
Fig. Reconstituída



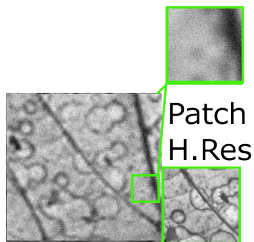
(a) z1: Cell3
Fig. Low Res.



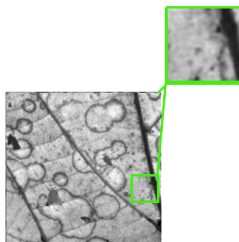
(b) Cell3
Fig. Original.



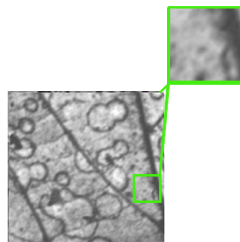
(c) x: Cell3
Fig. Reconstituída



(d) z1: Leaf
Fig. Low Res.



(e) Leaf
Fig. Original.



(f) x: Leaf
Fig. Reconstituída

Resultados Obtidos (Nossa Implementação)

Figura de Entrada	PSNR [dB]	ERRX	SSIM
CELL3	25,63	0,43	0,82
Leaf	48,31	0,62	0,87

Resultados do Autor

Figura de Entrada	PSNR [dB]	ERRX	SSIM
CELL3	21,99	0,82	0,76
Leaf	24,02	1,70	0,85




- Em teoria, minimizar as diferenças estruturais não locais e a diferença de histograma proporciona uma reconstrução textural mais eficiente.
- Ainda que os resultados sejam próximos, devido a correção necessária para tornar a implementação minimamente coerente, nada pode ser concluído em relação a falseabilidade do trabalho dos autores.
- Alguns pontos negativos: falta de referência para as imagens, os erros conceituais, a indisponibilidade de implementação autoral e a falta de retorno dos autores.
- Para garantir a falseabilidade, se deve reformular o algoritmo (e as derivadas) partindo do problema de otimização proposto pelos autores (tendo como base a referência [Condat 2013]).

- A implementação do cálculo de w_{ij} , cuja complexidade computacional varia com o tamanho da imagem estimada e o tamanho da amostra (*patch*) ($\mathcal{O}(NM)$), em MATLAB demanda um tempo muito grande em comparação com a implementação desse cálculo em C++/Eigen3. Por exemplo, de 6 horas para menos de 5 minutos (bom $\mathcal{O}(256^2 128^2)$). Outro ponto importante: a implementação em C++/Eigen3 utiliza vetorização SSE e AVX, que além de reduzir o tempo necessário quando não otimizado, melhora drasticamente a precisão numérica. De fato, no cálculo de w_{ij} têm-se uma função exponencial a qual é significativamente sensível a erros numéricos.

Tentativas de contato I

- Com Charles Deledalle, 3 de julho. Solicitando alguma implementação e informando que estava com dúvidas sobre como implementar as funções na forma como propostas no artigo. Sem resposta
- Com Mireille El Gheche, 7 de julho. Solicitando alguma implementação e informando que estava com dúvidas sobre como implementar o algoritmo. Sem resposta.
- Com todos os autores, 11 de julho. Informando as possíveis inconsistências e passando o acesso a nossa implementação disponível no GitHub (github.com/ysraell/texture_reconstruction). Sem resposta.

Obrigado pela vossa atenção.

-  CONDAT, L. A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms. *Journal of Optimization Theory and Applications*, v. 158, n. 2, p. 460–479, Aug 2013. ISSN 1573-2878. Disponível em: <http://dx.doi.org/10.1007/s10957-012-0245-9>.
-  GHECHE, M. E. et al. Texture synthesis guided by a low-resolution image. In: *2016 IEEE 12th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*. [S.l.: s.n.], 2016. p. 1–5.
-  GHECHE, M. E. et al. Texture reconstruction guided by a high-resolution patch. *IEEE Transactions on Image Processing*, v. 26, n. 2, p. 549–560, Feb 2017. ISSN 1057-7149.