Marouan Bouali, Alexander Ignatov - Image Processing Algorithm, Step-by-Step

Humna Sultan

Goal of Algorithm:

- “an adaptive algorithm was developed for operational use within the National Environmental Satellite, Data, and Information Service (NESDIS)’s SST system. The methodology uses a unidirectional quadratic variational model to extract stripe noise from the observed image prior to nonlocal filtering.”
- Remove thermal emissive bands that show residual striping → reduce the accuracy of measurements (SST, cloud masking, thermal front detection, etc.)

Algorithm Split Into 5 Parts

1. Noise Model: defining metrics of image noise
2. Directional Hierarchical Decomposition: decomposing noise functions
3. Nonlocal Filtering (NLM): removes noise while preserving image sharpness and detail
4. Optimization: derives a minimizer of the variational model used + Fourier transformations / space
5. Image Quality

Input: Striped image f

1: Initialize $\lambda_0 = 1$ and binary matrix $\mathbf{M}$

2: Solve $[u_0, v_0] = \inf\limits_{(u,v)/u+v=f} E_0(f, \lambda_0)$

3: While [NDF $(f, u_{k\geq 1}) < 0.95$]

Update $\lambda_k = \lambda_0/2^{k-1}$

Solve $[u_k, v_k] = \inf\limits_{(u,v)/u+v=v_{k-1}} E_k(v_{k-1}, \lambda_k)$

4: End

5: Apply nonlocal filter YNF to v_k

Ouput: Destriped image $= \sum_{i=0}^{k} u_i + \text{YNF}(v_k)$.

Mathematical Concepts:

- Partial derivatives + minimizing functionals
- Decomposition
- Infimum calculations
- Matrix applications
- Lagrange multipliers
- Element-by-element multiplication
- Exponential functions
- Fourier space / dimensions
- Energy functions
- Inverse and typical Fourier transformations

A: Noise Model

- images = bidimensional functions defined in Ω

★ Ω = bounded domain of $\mathbb{R}^2$

  - 2 dimensional $\rightarrow$ (x, y) where $(x, y \in \mathbb{R})$

$$f(x,y) = u(x,y) + \eta(x,y)$$

WHERE

- (x,y) = cartesian coordinates
- $f(x,y)$ = observed signal at pixel (x,y)
- $\eta(x,y)$ = stripe noise
- $u(x,y)$ = true signal to be estimated.

A: Noise Model (continued)

This equation solves for the observed signal at a provided pixel (x,y) by adding together the stripe noise associated with the provided pixel (x,y) and the true signal to be estimated associated with the provided pixel (x,y)

- in this case, η represents stripe noise $\rightarrow$ where geometrical considerations are more valuable than statistical assumptions.

B: Directional Hierarchical Decomposition

Definitions

* BV space = Space of functions of bounded variation.

- TNV approach: decompose an image in BV space into a sum of images defined in the intermediate (BV, L^2) space.
- method used in this paper: similar to TNV approach but restricted to the L^2 space and exploiting image gradient fields.

B: Directional Hierarchical Decomposition (continued)

Variational Decomposition

Variational decomposition:

$$[u_0, v_0] = \inf_{(u,v)/u+v=f} \underbrace{\int_\Omega \|\partial_x v\|^2 \, d\Omega + \lambda_0 \int_\Omega \|\partial_y (u - M.f)\|^2 \, d\Omega}_{\text{known as "energy functional"}}$$

- ★ ∂_x = partial derivative in the cross track direction
- ★ ∂_y = partial derivative in the across track direction
- ★ "." symbol = element-by-element multiplication.
- ★ $[u_0, v_0]$: represents a pair of functions
- ★ "inf" represents the infimum (greatest lower bound).
  - largest value that is less than / equal to every element in set

B: Directional Hierarchical Decomposition (continued)

Variational Decomposition (cont.)

- optimization is done over all pairs (u,v) such that $u+v=f$, where f is a provided function.

★ $\int_{\Omega} \|\partial_x v\|^2 \, d\Omega$ is the functional to be minimized

★ Ω is the domain of integration, aka the bounded domain of R^2

★ $\partial_x v$ = partial derivative of v with respect to x

B: Directional Hierarchical Decomposition (continued)

Variational Decomposition - explanation

Variational decomposition (continued):

* objective is to decompose a given function f into two parts: u and v
    * you want to minimize the smoothness of v, smoothness measured by $||\partial_x v||^2$

* this whole equation represents finding the decomposition (u_0, v_0) of f into u and v that minimizes the integral of the squared norm of the partial derivative of v over the domain Ω, subject to the constraint $u+v=f$

B: Directional Hierarchical Decomposition (continued)

Matrix M

Matrix M

$$M(x,y) = \begin{cases} 1 & \text{if } \{[\partial_x f(x,y)]^2 + [\partial_y f(x,y)]^2\}^{1/2} \geq \epsilon \\ 0 & \text{if } \{[\partial_x f(x,y)]^2 + [\partial_y f(x,y)]^2\}^{1/2} < \epsilon \end{cases}$$

- alleviates the problem of using the L^2 norm over the L^1 norm.
- M is defined using a threshold on the image gradient norm

★ for SST cases, $\epsilon = 0.8 \text{ K km}^{-1}$

B: Directional Hierarchical Decomposition (continued)

Matrix M (cont.)

★ the previous variational decomposition equation splits noisy observation f into...

- u_0 = initial estimate
- v_0 = striping and additional scale-dependent info.

★ $\partial_x f(x,y)$ = partial derivative of $f(x,y)$ with respect to x

★ $\partial_y f(x,y)$ = partial derivative of $f(x,y)$ with respect to y

B: Directional Hierarchical Decomposition (continued)

Extend Variational Decomp. With Lagrange

extend variational decomp. with Lagrange:

- Lagrange multiplier = strategy for finding the local maxima and minma of a function subject to equation constraints

$$[u_1, v_1] = \inf_{(u,v)/u+v=v_0} \int_\Omega \|\partial_x v\|^2 \, d\Omega$$

$$+ \lambda_0 \cdot 2^{-1} \int_\Omega \|\partial_y (u - M \cdot v_0)\|^2 \, d\Omega$$

* λ_0 = lagrange multiplier

* noisy observation = $f = u_0 + u_1 + v_0$.

B: Directional Hierarchical Decomposition (continued)

Extend Variational Decomp. With Lagrange (cont.)

* $\partial_x v$ = partial deriv. of v with respect to x
* $\partial_y (u - M.v_0)$ = partial deriv. of $(u - M.v_0)$ with respect to y
* optimization is done over all pairs (u,v) such that $u+v=v_0$, where v_0 represents striping and other scale-dependent info.
* "." = element-by-element multiplication.
* Ω = domain of integration + bounded domain of R^2

B: Directional Hierarchical Decomposition (continued)

Directional Hierarchical Decomp. derived from
Extended Variational Decomp.

$$[u_K, v_K] = \inf_{(u,v)\,/\,u+v=v_{K-1}} \int_\Omega \|\partial_x v\|^2 \, d\Omega$$

$$+ \lambda_0 \cdot 2^{-k} \int_\Omega \|\partial_y (u - M \cdot v_{K-1})\|^2 \, d\Omega$$

using this, f can also be expressed now as...

$$\left.\begin{aligned} f &= u_0 + v_0 \\ &= u_0 + u_1 + v_1 \\ &= \\ &\;\vdots \end{aligned}\right\} = \sum_{i=0}^{K} u_i + v_K .$$

B: Directional Hierarchical Decomposition (continued)

Directional Hierarchical Decomp. derived from Extended Variational Decomp. (cont.)

- DHD acts as a directional filter in the spacial domain that progressively retrieves cross-track variations in the term v_K while isolating the stripe noise in its high-frequency domain.

- We know that as the number of iterations approaches ∞, the Lagrange multiplier $\lambda_0 \cdot 2^{-K}$ converges to 0

$$\Downarrow$$

$[u_K, v_K]$ converge into one sol. of the problem:

$$[u_K, v_K] = \inf_{(u,v)/u+v=v_{K-1}} \int_\Omega \| \partial_x (u - v_{K-1}) \|^2 \, d\Omega$$

B: Directional Hierarchical Decomposition (continued)

Directional Hierarchical Decomp. derived from Extended Variational Decomp. (cont.)

Knowing that

$u_K = v_K + A$

$\forall (x,y) \in \mathbb{R}^2, \partial_x A(x,y) = 0$

★ $\forall$ = for all

★ A = matrix that represents one constant per line.

for all (x,y) that is an element of $\mathbb{R}^2$ (dim. 2),
the partial derivative of $A(x,y)$ with respect to x is 0.

C: Nonlocal Filtering

Weighting Function that measures the Radiometric Similarity Between Pixels (x, y) and (x’, y’)

$$w(x,y,x',y') = \exp\left\{\frac{-[v_N(x,y) - v_N(x',y')]^2}{2\sigma^2}\right\}$$

★ σ controls the decay of weighting coefficients with respect to the radiometric distance between pixels + depends on the NEΔT.

★ NEΔT = noise equivalent delta temperature – represents the radiometric resolution and sensitivity of a radiometer.

★ σ^2 = NEΔT/2

val. of noisy pixel

★ VN = remaining information, also contains a low-freq. component that has to be retrieved for the estimation of the true scene.

C: Nonlocal Filtering (continued)

Yaroslavsky Neighborhood Filter (YNF)

Yaroslavsky neighborhood filter (YNF).

$$YNF[v_N(x,y)] = \frac{1}{C(x,y)} \int_{N(x,y)} v_N(x,y)\, w(x,y,x',y')\, dx'\, dy'$$

★ $C(x,y)$ = normalization term represented by...

$$C(x,y) = \int_{N(x,y)} w(x,y,x',y')\, dx'\, dy'$$

★ $N(x,y)$ = spatial neighborhood of pixel (x,y)

C: Nonlocal Filtering (continued)

Defining N(x,y) in YNF

$$N(x,y) = \{(x', y') \in \Omega' : |y-y'| \le \Delta y \text{ and } |x-x'| = 0\}$$

★ Ω' = subdomain of Ω defined with respect to the binary matrix M as...

$$\Omega' = \{(x,y) \in \Omega : M(x,y) = 0\}.$$

★assuming an even # of detectors, $\Delta y = D/2$

★D = # of detectors. (16 for VIIRS "M" bands)

★to account for possible mirror side differences, $\Delta y = D$ (no mirror banding in VIIRS banding).

C: Nonlocal Filtering (continued)

Nonlocal filtering Yaroslavsky Neighborhood Filtering (YNF)

$$YNF[v_N(x,y)] = \frac{1}{C(x,y)} \int_{y-\Delta x}^{y+\Delta x} v_N(x,y') \exp\left(\frac{-B^2}{2\sigma^2}\right) dy'$$

* exp = exponential function ().
* $B = v_N(x,y) - v_N(x,y')$

- for N iterations of PHD, an estimate of true signal is...

$$\hat{u} = \sum_{K=0}^{N} u_K + YNF[v_N(x,y)].$$

C: Nonlocal Filtering (continued)

Nonlocal filtering Yaroslavsky Neighborhood Filtering (YNF) - (cont.)

meaning that the estimate of true signal is equal to the summation of all u values from $k=0$ to N added to the YNF of provided value of noisy pixel (x,y).

D: Optimization

Redeclaring Energy Functional:

redeclaring energy functional:

$$\begin{aligned} E_0(f, \lambda_0) = & \int_\Omega \|\partial_x (u-f)\|^2 \, d\Omega \\ & \underbrace{+ \lambda_0 \int_\Omega \|\partial_y (u - M \cdot f) \, d\Omega}_{\text{this satisfies the below Euler-Lagrange eq.}} \end{aligned}$$

$$\frac{\partial E_0}{\partial f} - \frac{\partial}{\partial x}\left[\frac{\partial E_0}{\partial(\partial_x f)}\right] - \lambda_0 \frac{\partial}{\partial y}\left[\frac{\partial E_0}{\partial(\partial_y f)}\right] = 0$$

D: Optimization (continued)

Redeclaring Energy Functional (cont.) - defining vars

- ☆ E_0 = energy functional
- ☆ f = noisy observation.
- ☆ x and y represents a pixel (x, y)
- ☆ λ_0 = lagrange multiplier previously defined
- this equation simplifies to...

$$\partial_x^2 (u-f) + \lambda_0 \partial_y^2 (u - M \cdot f) = 0$$

- ☆ u is derived from (u_0, v_0) in original energy functional
- the 2nd partial derivative of $(u-f)$ with respect to x plus the 2nd partial derivative of $(u - M \cdot f)$ with respect to y multiplied by the lagrange multiplier = 0.

D: Optimization (continued)

Spatial Frequency + Fourier Domain

★ (ξ_x, ξ_y) = spatial frequency vars in Fourier domain

- for a Fourier domain image, each point represents a particular frequency contained in the spatial domain image.

★ i = imaginary unity

★ $\mathcal{F}$ = Fourier transform of a derivative function.

★ f = noisy observation,

$\mathcal{F}(\partial_x f) = i \cdot \xi_x \mathcal{F}(f)$ and $\mathcal{F}(\partial_y f) = i \cdot \xi_y \mathcal{F}(f)$

$\mathcal{F}(\partial_x^2 f) = -\xi_x^2 \mathcal{F}(f)$ and $\mathcal{F}(\partial_y^2 f) = -\xi_y^2 \mathcal{F}(f)$.

D: Optimization (continued)

Fourier Transformation Applied to Simplified Euler-Lagrange Equation Derived From Energy Functional

Fourier transform applied to simplified Euler-Lagrange equation derived from energy functional:

★ $\tilde{a}$ = Fourier transform of a

★ a^M = element-by-element multiplication of a with matrix M

$$\tilde{u}_0 = \left(\xi_x^2 \tilde{f} + \lambda_0 \xi_y^2 \widetilde{f^M}\right)\left(\xi_x^2 + \lambda_0 \xi_y^2\right)^{-1}$$

$$\tilde{v}_0 = \left[\lambda_0 \xi_y^2 \left(\tilde{f} - \widetilde{f^M}\right)\right]\left(\xi_x^2 + \lambda_0 \xi_y^2\right)^{-1}$$

- inverse Fourier transform $\rightarrow$ estimate of $[u_0, v_0]$ in spatial domain

D: Optimization (continued)

Minimizing Energy Functional at Nth Iteration

minimizing energy functional $E_N(v_{N-1}, \lambda_N)$ at Nth iteration

$$E_N(v_{N-1}, \lambda_N) = \int_\Omega \|\partial_x (u - v_{N-1})\|^2 \, d\Omega$$

$$+ \lambda_0 \cdot 2^{-N} \int_\Omega \|\partial_y (u - M \cdot v_{N-1})\|^2 \, d\Omega$$

★ λ_N = Lagrange multiplier associated with N

★ v_N = val. of noisy pixel

★ Ω = bounded domain of R^2.

D: Optimization (continued)

Simplify Euler-Lagrange Equation

Simplify Euler-Lagrange equation:

$$\partial_x^2 (u - v_{N-1}) + \lambda_0 \cdot 2^{-N} \partial_y^2 (u - M \cdot v_{N-1}) = 0$$

- we understand all the variables

D: Optimization (continued)

Shifting Euler-Lagrange to Fourier Space

$$-\xi_x^2(\widetilde{v_{N-1}} - \tilde{u}) - \lambda_0 . 2^{-N} \xi_y^2 (\widetilde{v^M_{N-1}} - \tilde{u}) = 0$$

shifting to Fourier space

define Nth iterations in Fourier Space:

$$\tilde{u}_N = (\xi_x^2 \widetilde{v_{K-1}} + \lambda_0 \cdot 2^{-k} \xi_y^2 \widetilde{v^M_{K-1}})(\xi_x^2 + \lambda_0 \cdot 2^{-k} \xi_y^2)^{-1}$$

$$\tilde{v}_N = [\lambda_0 \cdot 2^{-k} \xi_y^2 (\widetilde{v_{K-1} - v^M_{K-1}})](\xi_x^2 + \lambda_0 . 2^{-k} \xi_y^2)^{-1}$$

E: Image Quality

NIF and NDF Definitions

NIF and NDF

★**NIF** = metric that quantifies relative change in gradients across the scan

★**NDF** = quantifies the relative change with destriping in the gradients along the scan direction

★ Ω' = subdomain of Ω defined with respect to the binary matrix M as...

$$\Omega' = \{(x,y) \in \Omega : M(x,y) = 0\}$$

E: Image Quality

NIF and NDF Applications

$$NIF(f,u) = \left(\int_{\Omega'} \left(|\partial_y f| - |\partial_y u| \right) dy \right) \left(\int_{\Omega'} |\partial_y f| \, dy \right)^{-1}$$

$$NDF(f,u) = 1 - \left[\int_{\Omega'} |\partial_x (f-u)| \, dy \right] \left(\int_{\Omega'} |\partial_x f| \, dx \right)^{-1}$$

Concluding Notes

- the algorithm and method described in this paper provides an improvement in stripe noise, not particularly a cosmetic improvement such as suggested in other algorithms
- Algorithm utilizes a series of equations, simplifications, generalizations, and transformations
- Iteration is critical for Step 3 (slide 4)
- This algorithm is altered in other studies (ifor ex: “Destriping algorithm for improved satellite-derived ocean color product imagery”)