Two-dimensional Phase Unwrapping Method Using Cost Function of L^0 Norm

J GAO [1,2], L LI[3] and L SHI[3]

[1] College of geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
[2] Health Services Network Systems Engineering Research Centre of Education Ministry, Nanjing, 210023 China
[3] College of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, 210023, China

E-mail: gaoj@njupt.edu.cn

Abstract. Considering cost model and convergence speed of the minimum norm unwrapping, a highly efficient two-dimensional global phase unwrapping method optimized with L^0 norm is proposed. As analysing features of cost model in phase unwrapping with minimum norm, a cost function definition is provided in line with the L^0 norm, which impose a stronger constraint in the tangent direction of phase discontinuity boundary than that in normal direction, in order to preserve integrity of discontinuity during iterative unwrapping processing for continuous phase. For the sake of slow speed of low-frequency error convergence during linear solving, a data partitioning strategy is introduced into unwrapping processing. Due to independence of minimum norm method in blocks, linear solving only focus on high-frequency information and improve efficiency of iterative work, and the low-frequency processing part is transferred to offsetting-aligning between blocks. With experiments and analysis, reliability and efficiency of the novel phase unwrapping method are certified comparing to existing methods.

1. Introduction

Interferometric Synthetic Aperture Radar or Sonar (InSAR/InSAS) has arisen in topographic observing for decades. Wherein, two-dimensional phase unwrapping is a key step of DEM generating in interferometry. Existing unwrapping algorithms can be basically divided into two classes, path following and global optimization.

Unwrapped phase is obtained by integrating phase gradients along paths from reference point in following algorithms, auxiliary information from wrapped phase or external data are taken into account to avoid disturbance of phase discontinuity by posing barriers or setting priority levels to integrating paths. Such algorithms include branch-cuts method [1], the quality map guided method [2-4], Flynn minimal discontinuity method [5], network flow method [6-7], and these belong to local methods. Global optimization unwrapping is mostly an approaching towards the global extremal value of optimization objective function in order to obtain the optimal solution. The global methods usually take cost functions in form of norms, also known as minimum norm methods, including least-squares method, L^p Norm method [8].

Minimum norm unwrapping method take measure of difference between unwrapped and wrapped phases on gradient as cost function, and search for the optimal solution by minimizing the objective function globally. One classical approach is taking least squares method with L^2 norm, which translates unwrapping problem into a Poisson's equation, but it is unpractical for the sake of detail loss by reason of over-smoothing. On basis of the least square method, weighted least squares method [8] was presented to improved conditions. Ghiglia [8] introduced a general framework L^p norm method, as a generalization of minimum norm unwrapping methods. Recently a variety of mixed norm [9] was proposed in unwrapping program.

Minimum norm unwrapping research focus on selecting an appropriate cost model, another application difficulty is low convergence speed of iterative procedure. Based on existing achievement, a cost function similar to L^0 norm is provided here for two-dimensional phase unwrapping. A data partitioning scheme is introduced to overcome the low-frequency error convergence problem in linear differential equation solving, and improve the efficiency of unwrapping process.

2. Principle

2.1. Norms in phase unwrapping

The relation between wrapped phase ψ and true unwrapped ones φ can be described like this

$$\psi = W(\varphi)$$

Where W is the wrapping function modulo 2π. Phase unwrapping is a process of recovering φ from ψ, and when a result meets certain constraint from true values, that means the unwrapping process achieves a desired goal. The unwrapping resolution guidelines generally are about optimization issues of gradient differences, such as

$$\inf_{\phi} \int_{\Omega} g(|\nabla u|)\,\mathrm{d}x \tag{1}$$

Where $u = \phi - \varphi$, ϕ is the unwrapped phase, Ω refers to definition domain of phase, $g(\cdot)$ is so-called cost function making effects of derivative difference in the statistics. Such optimization constraint makes unwrapped result ϕ maintain some consistency with φ.

Assuming $g(\cdot)$ is a continuous function and first order derivable, the corresponding Euler-Lagrange equation of minimal problem in Eq. (1) is

$$\mathrm{div}\left(\frac{g'(|\nabla u|)}{|\nabla u|} |\nabla u| \right) = 0 \tag{2}$$

Where div is the divergence operator. In the condition of Nyquist theorem, gradient of u can be obtained by

$$\nabla u = \nabla \phi - \nabla \varphi = \nabla \phi - \nabla W(\psi) \tag{3}$$

A key of building Eq. (2) is definition of the cost function g, which plays a dominant role in unwrapping process. The primary difficulty of unwrapping just lies in local phase discontinuity caused by distinct features of object observed or environmental noises, that contradicts with given assumptions in Eq. (3). If phase discontinuities occur, cost function should preserve congruency of continuous data, and meanwhile suppress disturbance from phase discontinuity.

Different definitions of the cost function would bring about different results, for example, in form of L^0, L^1 and L^2 norms which are shown in (a) of Fig. 1 [10]. As one-dimensional example shown in (b), there are two fixed values at points A with zero phase residual u and B with nonzero residual. Among the three given connection paths, curve l_1 refers to an arbitrary monotone function, curve l_2 refers to a line with constant derivative value and curve l_3 refers to a connection with minimal

residual influence. And curve l_2 is accordant minimal L^2 norm of ∇u , l_3 corresponds to L^0 norm and all the three curves to L^1 norm.

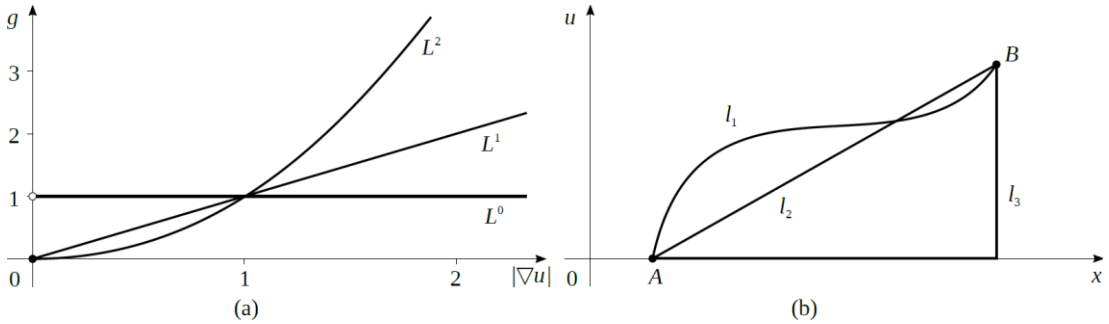

Figure 1. Cost function in form of L^0, L^1 and L^2 norms (a) and residual of different connection paths (b) between point *A* and *B*.

The best choice is L^0 norm cost function [10], while L^2 norm take strong smoothing effects, and L^1 norm is prone to staircase phenomenon. But classic L^0 norm function is not continuous or first order derivable, and it is a typically and untreatable problem (NP Hard) in network-flow optimization. Here a novel expression of L^0 norm cost function is proposed for applicable minimal norm unwrapping method.

2.2. L^0 norm cost function

In the two-dimensional case, the divergence in Eq. (2) can be decomposed into two second order derivate parts, $c_\xi u_{\xi\xi} + c_\eta u_{\eta\eta}$, where $\eta = \frac{\nabla u}{|\nabla u|}$ and ξ the normal vector to η , $u_{\xi\xi}$ and $u_{\eta\eta}$ refer to second order derivate in ξ and η directions. Namely,

$$\operatorname{div}\left(\frac{g'(|\nabla u|)}{|\nabla u|}|\nabla u|\right) = \nabla\left(\frac{g'(|\nabla u|)}{|\nabla u|}\right)\cdot\nabla u + \frac{g'(|\nabla u|)}{|\nabla u|}\Delta u = \frac{g'(|\nabla u|)}{|\nabla u|}u_{\xi\xi} + g''(|\nabla u|)u_{\eta\eta}$$

Where Δ is Laplacian operator and

$$u_{\xi\xi} = \frac{1}{|\nabla u|}(u_x^2 u_{yy} + u_y^2 u_{xx} - 2u_x u_y u_{xy}) \quad u_{\eta\eta} = \frac{1}{|\nabla u|}(u_x^2 u_{xx} + u_y^2 u_{yy} + 2u_x u_y u_{xy})$$

Equation (2) can be rewritten as [11]

$$\frac{g'(|\nabla u|)}{|\nabla u|}u_{\xi\xi} + g''(|\nabla u|)u_{\eta\eta} = 0 \tag{4}$$

Such a second order differential equation can be taken as a smoothing constraint, and coefficient values of derivate modulate the acting strength. Normally, smoothing in ξ and η directions both are expected to work in continuous phase region, whereas only in ξ is expected to run at discontinuity phase. That means

$$\lim_{s\to 0^+}\frac{g'(s)}{s} = \lim_{s\to 0^+} g''(s) = g''(0) > 0 \quad \lim_{s\to+\infty}\frac{g'(s)}{s} > 0 \quad \lim_{s\to+\infty} g''(s) = 0 \tag{5}$$

But the two conditions are not compatible, Equation (5) can be moderated as following

$$\lim_{s\to 0^+}\frac{g'(s)}{s} = \lim_{s\to 0^+} g''(s) = g''(0) > 0 \quad \lim_{s\to+\infty}\frac{g'(s)}{s} = 0 \quad \lim_{s\to+\infty}\frac{sg''(s)}{g'(s)} = 0 \tag{6}$$

Then gradient coefficients of the two components approach to zero in this case, and the tangent component more slowly than the normal component parts. In fact, domain of variable *s* is usually finite, namely, no more than a half cycle in wrapped phase data.

According to requirements from (6) for building cost function, Geman-Reynolds model [12] scheme is employed here for L^0 norm, namely

$$g(|\nabla u|) = \frac{|\nabla u|^2}{\kappa^2 + |\nabla u|^2} \tag{7}$$

Where κ is a positive scale coefficient. The function value at $|\nabla u| = 0$ is zero, and tends to a constant when $|\nabla u|$ get large, the cost function with different scale coefficient is shown in (a) of Fig. 2. And the corresponding dominating smoothing coefficient c_ξ is shown in (b). If value of κ is too large, for example $\kappa = 1$, the cost function approximate to L^1 norm, and if too little, effect of the coefficient c_ξ is limited in a small range, just like that with $\kappa = 0.01$ in (b) of Fig. 2. Empirical value of κ is about 0.1.

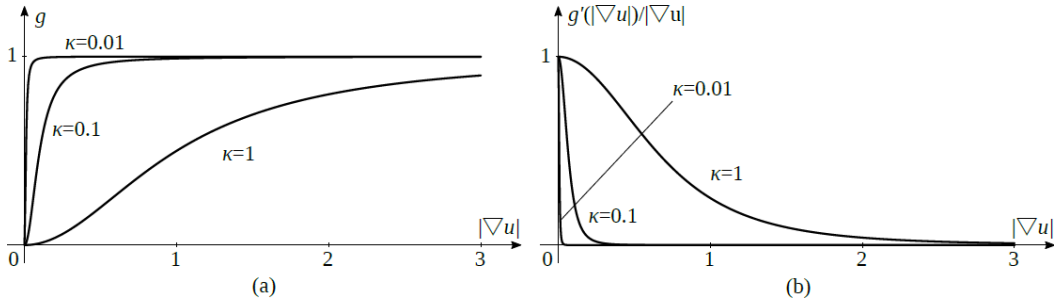

Figure 2. Cost function (a) and coefficient c_ξ(b) with different scale coefficients.

Substituting the cost function (7) to (2), a new unwrapping equation is obtained

$$\operatorname{div}\left(\frac{\nabla u}{(\kappa^2 + |\nabla u|^2)^2}\right) = 0 \tag{8}$$

Which describes constraint of L^0 norm unwrapping problem and is not easy to solve directly.

3. Numerical method

3.1. Solution of nonlinear equation

Dealing with numerical differential equation with discrete data, the first thing need to do is converting differential equations into difference ones, and then make them solved. An iterative scheme is employed here, each iteration runs on the basis of its predecessor during the cycle procedure, and the difference between results of the two-step process become an important part to focus on.

According to Eq. (3), we let

$$b = 1/(\kappa^2 + |\nabla\phi - \nabla W(\varphi)|^2)^2$$

Which can be taken as a constant, when solving

$$\operatorname{div}(b\nabla\phi) = \operatorname{div}(bW(\nabla\varphi)) \tag{9}$$

And its solution would be used to update *b*. Assuming unwrapping results of step *n* and step *n+1* meet with

$$\operatorname{div}(b_n\nabla\phi_{n+1}) = \operatorname{div}(b_n W(\nabla\varphi))$$

By means of correction value, it can be rewritten as

$$\operatorname{div}(b_n\nabla\delta\phi_n) = \operatorname{div}(b_n(W(\nabla\varphi) - \nabla\phi_n))$$

Where $\delta\phi_n = \phi_{n+1} - \phi_n$. In this case the right-hand-side refers to original wrapped phase information exclude known unwrapped part, then solving for correction values $\delta\phi_n$ and a new unwrapped phase in loops for final result. Such a linear equation with sparse coefficient matrix, whose condition number is often large, and the solver generally take Krysolv subspace algorithms, just like Preconditioned Conjugate Gradient (PCG) method [13-14] employed here.

The cycle terminates in some conditions, such as inexistence of residual point in residual wrapped phase $\delta\psi$

$$\delta\psi_n = \psi - \phi_n$$

When there is no phase residue, it suggests no phase discontinuity in $\delta\psi_n$, whose unwrapped counterpart can be calculated following arbitrary integration paths. The whole process of unwrapping is the following.

1. let n = 0, initialize unwrapped phase by $\phi_n = 0$;
2. calculate coefficient matrices b_n;
3. calculate right-hand-sides of equations $\mathrm{div}\left(b_n(W(\nabla\varphi)-\nabla\phi_n)\right)$;
4. solve equations to obtain solutions $\delta\phi_n$;
5. update unwrapped phase $\phi_{n+1} = \phi_n + \delta\phi_n$;
6. calculate residual wrapped phase $\delta\psi_{n+1}$;
7. check existence of residue, if true, let n=n+1 and go to step 2;
8. unwrap residual unwrapped phase $\delta\psi_{n+1}$, finish unwrapping process.

3.2. Data partitioning

For the sake of low-frequency residual attenuation, if the size of wrapped phase is very large, inefficiency in convergence of linear equations will be more evident. Such a time-consuming method would not be practical in real interferometry processing.

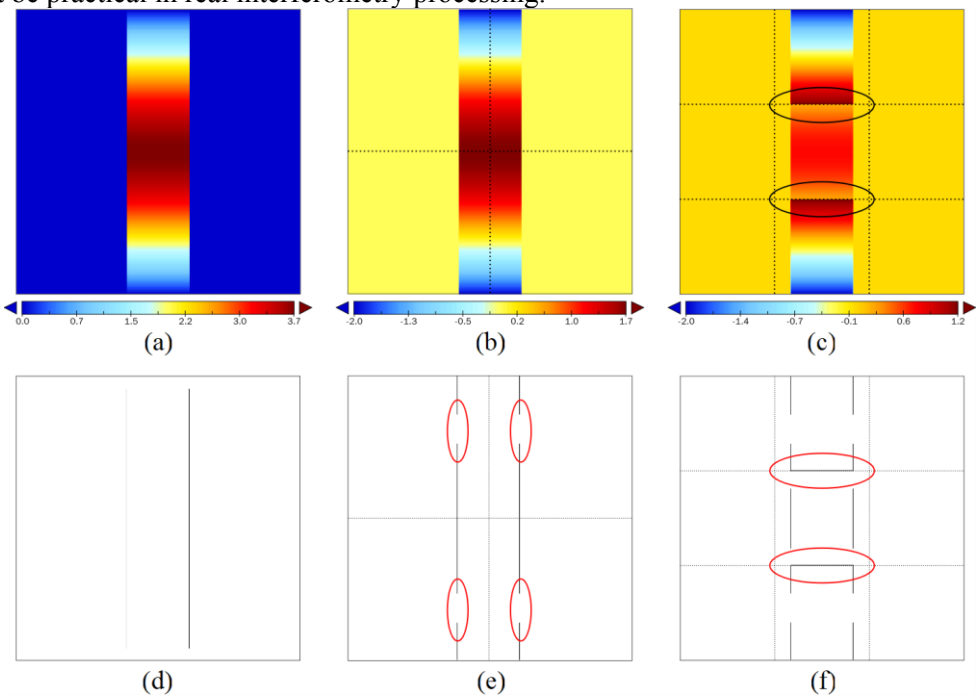

Figure 3. Unwrapping results and discontinuity edges of original, 2×2 blocks and 3×3 blocks.

Taking into account of data partitioning, big blocks are divided into small size ones, then unwrapping is carried out on small size data. After unwrapping, small blocks would be adjoined together for an integral result. Only with high frequency residual, solving on small block data would convergence faster. Parallel processing mode can be adopted in unwrapping of small blocks for their independence.

The only thing need to note here is that, partitioning should not divide discontinuity boundaries into little chips, when discontinuity of wide range appears, though that occur scarcely. For example,

synthetic phase data is generated by adding a vertical half-cycle sine band to background plane, shown as (a) of Fig. 3 and the discontinuity boundaries are shown in (d).

The synthetic phase data, is divided into 2×2 and 3×3 blocks respectively. After unwrapping on each sub-block and then assembling for the unwrapping results. The results and corresponding phase discontinuity boundaries are shown in (b), (c), (e) and (f). As in the former scheme, discontinuity boundaries are broken, though that does not destroy the unwrapping result.

In the latter scheme, the discontinuity boundaries are divided into more little sections. The smaller blocks scheme prevents to recover the wrapped phase correctly, there is an error in the central location.

Therefore, when designing a data partitioning scheme, it needs attention to choose block size, avoid breaking discontinuity boundaries, although process run faster on smaller blocks. In general, the phase discontinuity is often due to signal noise or abnormal feature of target observed, and there is seldom a discontinuity boundary over wide range, so this problem is not very serious.

4. Experiments and analysis

The machine used for processing in experiments is equip a CPU with 8 logical cores at 2.7GHz, dual-channel 8GB DDR3 memories at 1600MHz and a 64-bit operating systems. Processes work in a shared memory multi-core parallel processing mode [15], using a linear equation solver of good stability, half coarsening multi-grid preconditioned conjugate gradient method.

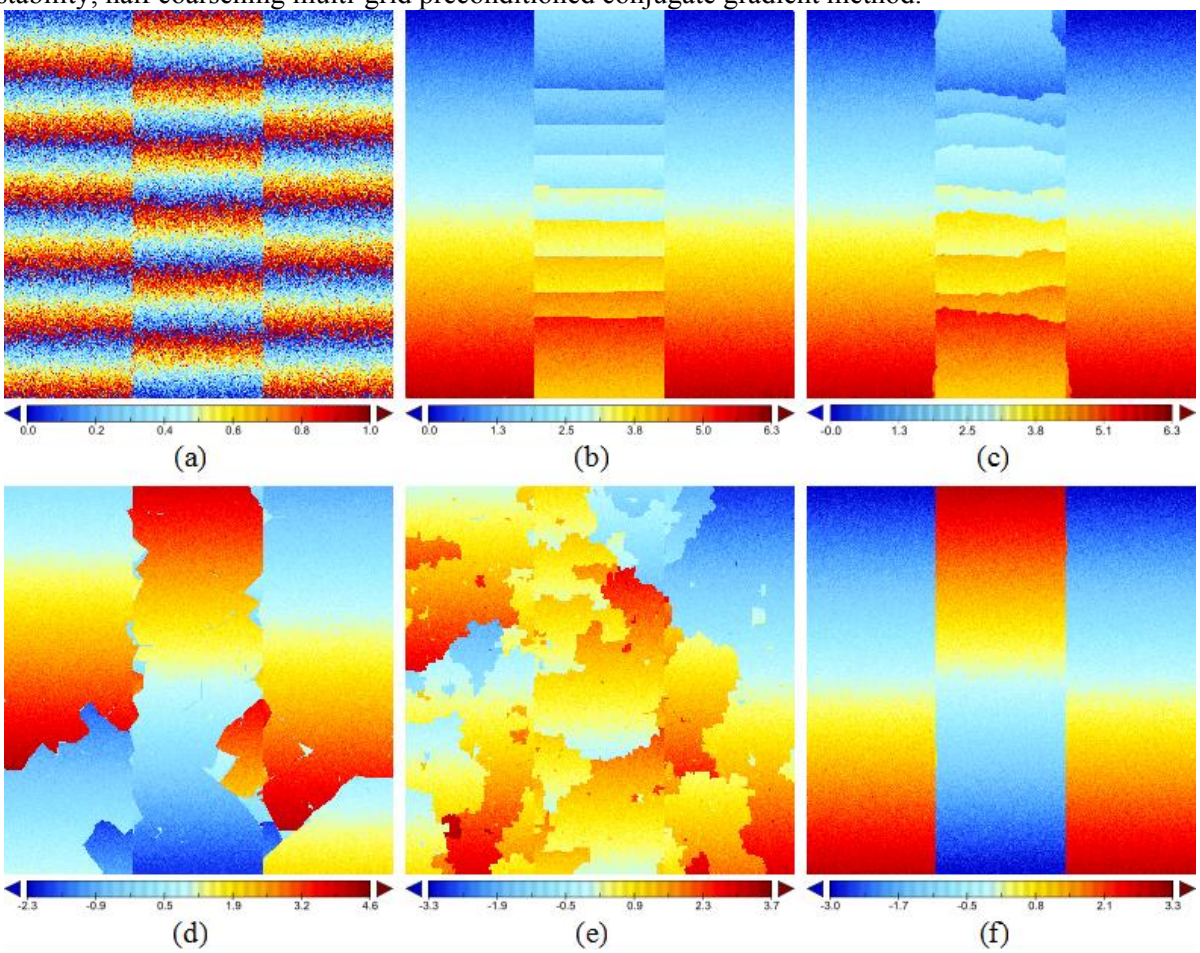

Figure 4. Wrapped noisy shear phase data (a) and unwrapped phase from Flynn's (b), MCF method(c), Goldstein's method (d), quality-guided method (e) and proposed method (f).

Experimental data include synthetic wrapped phase data and real InSAR data, the synthetic ones are generated by wrapping noisy shear data, 257×257 pixels, shown as (a) of Fig. 4. The synthetic data are unwrapped by Flynn's Minimal Discontinuity method, Minimum Cost Flow (MCF) method,

Goldstein's branch cut method, derivative variance based Quality Map Guided method and the proposed method. The results are shown in Fig. 4.

As shown, Flynn's and MCF and methods failed in recovering at central area for their L^1 norm cost schemes, Goldstein's branch cut method with L^0 norm get a better result but not eligible, for wrong connections of residues. The Quality Map Guided method did not well for fragment data and the proposed L^0 norm method works well.

Using real InSAR interferometric phase as the experimental data, whose size is 1398×622 pixels, contains a global gently fluctuation, and serious noise at several locations. Especially some noises are located in areas near boundaries, whose bad impact would bring obstacle to convergence for stretching from only one side, and cycle operation within a larger area may be more time-consuming.

Experimental data are divided into 7×3 and 14×6 blocks at 200-pixel and 100-pixel scales respectively and then unwrapping. As shown in Fig. 5, the original InSAR wrapped phase data (a), 14×6 partitioning diagram (b), partition unwrapping results (c), offsets of blocks (d), final results (e), the corresponding discontinuity edges (f), it suggests that unwrapping the result is able to maintain good integrity even partitioned in 100-pixel scale.

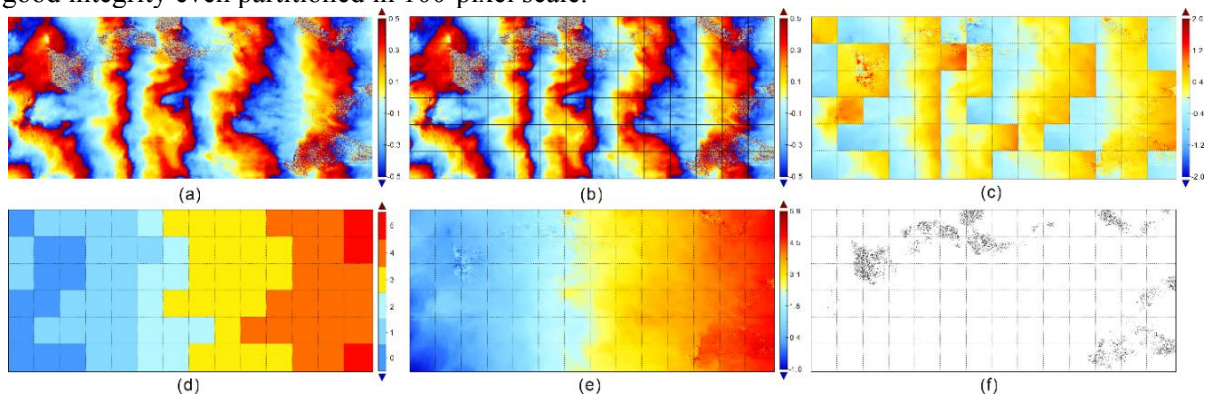

Figure 5. Unwrapping process of InSAR phase in 14×6 blocks.

In order to verify validity of proposed method, different unwrapping algorithms run on the InSAR phase data. The quality of unwrapping a result and the efficiency of unwrapping process are described by quantitative statistical indicators, including standard deviation of difference between re-wrapped results and original data, length of phase discontinuity edges, and the running time of process, statistical values are shown in Tab. 1.

It should be noted that the data are stored in 8-bit float type, and there are no more than 7 significance digits for a value in such case, so if the order of a precision index is higher than 10^{-6}, that is mainly caused by truncation errors. As shown in the table, the standard deviation of Goldstein branch cut method is more prominent, mainly because the method may bring about islands or hole in the noisy area of poor quality. Considering indicators of length of discontinuity boundaries, Flynn's is best, whose most important principle just is that, minimal discontinuity. The proposed method, in both partitioning schemes get good performance better than Goldstein's, Quality Map Guided and SNAPHU.

Table 1. Statistics Indicators of the Results from Different Unwrapping Methods

Algorithm	L^0 norm/px	L^1 norm/px	Time/sec
Goldstein's	17546	34553	0.156
Quality-Guided	62881	124229	3.927
Flynn's	11984	11989	6.938
MCF	12997	12997	15.265
SNAPHU	12989	12989	3.534
PUMA	13012	13012	152.507
Proposed(7×3)	13714	13730	17.979
Proposed(14×6)	13729	13748	7.564

Finally, taking into account of running time, the branch-cut is fastest due to its simplicity on operation, followed by the quality-guided and PUMA slowest method. The running time of proposed method in this paper, in both partitioning scheme, are all better than SNAPHU's and PUMA's, and especially in scheme of 14×6 blocks which is also better slightly than Flynn's. That proves the efficiency of new method.

Owing to the table, the proposed unwrapping method achieve a good performance in experiments, obtaining reliable result with small discontinuity and big efficiency. That are not found in classical minimal norm unwrapping methods.

5. Conclusion

This paper proposed a global phase unwrapping method using cost function of L0 norm, which makes different constraints intensity on normal and tangent directions at discontinuity edges, obtaining reliable unwrapped phase. The strategy of data partitioning gets rid of low-frequency residual in linear differential equation to improve processing efficiency. The experimental results and comparing with other popular methods verify its validity and efficiency.

Acknowledgements

This work was supported by the Scientific Research Foundation of Nanjing University of Posts and Telecommunications under Grant NY213056, NY215109, by the Natural Science Foundation of Jiangsu Province under Grant BK20130864, and by the National Natural Science Foundation of China under Grant 41401480, 41501378, 41501497.

References

[1] Huang Q., 2015. Parallel branch-cut algorithm based on simulated annealing for large-scale phase unwrapping. *IEEE Transactions on Geoscience and Remote Sensing*, 53(7), pp.3833–3846.

[2] Gao D., 2012. Mask cut optimization in two-dimensional phase unwrapping. *IEEE Geoscience and Remote Sensing Letters*, 9(3), pp.338–342.

[3] Zhong H., 2014. A quality-guided and local minimum discontinuity based phase unwrapping algorithm for InSAR/InSAS interferograms. *IEEE Geoscience and Remote Sensing Letters*, 11(1), pp.215–219.

[4] Liu G., 2014. A new quality map for 2-d phase unwrapping based on gray level co-occurrence matrix. *IEEE Geoscience and Remote Sensing Letters*, 11(2), pp.444–448.

[5] Flynn T., 1998. Phase unwrapping using discontinuity optimization. In: *Geoscience and Remote Sensing Symposium Proceedings*, Vol. I, pp.80–82.

[6] McKilliam R. G., 2014. Polynomial phase estimation by least squares phase unwrapping. *IEEE Transactions on Signal Processing*, 62(8), pp.1962–1975.

[7] Dong J., 2015. Simultaneous phase unwrapping and removal of chemical shift (spurs) using graph cuts: Application in quantitative susceptibility mapping. *IEEE Transactions on Medical Imaging*, 34(2), pp.531–540.

[8] Ghiglia D. C., 1994. Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods. *Journal of the Optical Society of America*, A 11(1), pp.107-117.

[9] Liu H., 2015. A novel mixed-norm multi-baseline phase-unwrapping algorithm based on linear programming. *IEEE Geoscience and Remote Sensing Letters*, 12(5), pp.1086–1090.

[10] Chen C. W., 2001. Statistical-cost Network-flow Approaches to Two-dimensional Phase Unwrapping for Radar Interferometry. Ph.D. thesis, Stanford University.

[11] Kornprobst P., 1997. Nonlinear operators in image restoration. In: *IEEE Comput. Soc.*, pp.325-330.

[12] Geman D., 1992. Constrained restoration and the recovery of discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(3), pp.367–383

[13] Trefethen L. N., 1997. Numerical Linear Algebra, Society for Industrial and Applied Mathematics.
[14] Demmel J. W., 1997. Applied Numerical Linear Algebra, SIAM.
[15] Barbara C., 2007. Using OpenMP: Portable Shared Memory Parallel Programming. The MIT Press.