TURBO ITERATIVE SIGNAL PROCESSING

Hong SUN[1,2], and Henri Maître[2]

[1]Wuhan University, Wuhan; [2]Telecom-ParisTech, Paris

ABSTRACT

A Turbo iterative method for signal processing is proposed. This method is a kind of multi-systems collaborative signal processing through iteration: several independent systems work in rotation, and each system takes feedback information from the other systems as a priori condition. We have applied such a Turbo iterative signal processing (TISP) method on speech signal enhancement, and on SAR (synthetic aperture radar) image filtering, segmentation and fusion. Some practical results presented in this article show that the Turbo iterative algorithm converges after 5-10 iterations and it improve greatly the signal processing performance. The TISP also shows an effective machine-learning method, that is making a discussion between several independent systems through Turbo iteration.

***Index Terms*—** Iterative signal processing method, Turbo iteration, machine learning, image processing, speech processing

1. INTRODUCTION

Turbo Code was proposed by Berrou et al. in 1993 [1], and it is attention-getting in the field of coding theory for its performance rather impressive. We have seen that, Turbo code not only is an excellent error-correcting code for a very noisy channel, but also it tells us a collaborative signal processing approach in complex cases.

We proposed to employ the Turbo iterative approach to synthetic aperture radar (SAR) image processing in 2000 [2]. And in these years, we are successful in applying the Turbo iterative method on SAR image filtering [3], segmentation [4], fusion [5], feature extraction [6] and speech signal enhancement [7], improving greatly the performance of signal processing.

A unified principle of Turbo iterative signal processing (TISP) method is presented in this article, and the effectiveness of this method is demonstrated by several examples of its applications on image processing and speech processing. In addition, the TISP method reveals us a kind of a machine learning approach: several independent systems discuss together and learn from each other, so that the performance will be greatly improved.

The remainder of this article is organized as follows. In Section 2, the basic principle of Turbo iterative signal processing is presented; In Section 3 and Section 4, the principle of SAR image filtering and of speech enhancement based on Turbo iteration method is discussed and some results are shown; In Section 5, some techniques of Turbo iterative image processing, including image segmentation and fusion, are briefly studied; Finally, section 6 is our concluding discussion.

2. PRINCIPLE OF TURBO ITERATIVE SIGNAL PROCESSING

2.1. Turbo Iterative Signal Processing (TISP)

Turbo iterative signal processing (TISP) is a multi-systems collaborative working method. We describe this method by a block diagram in Fig.1 with two systems without loss of generality.

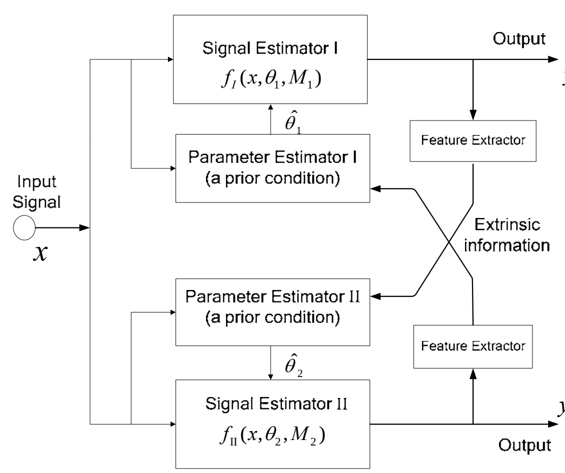


Fig.1. Turbo iteration method of signal processing

All systems (a “signal estimator” in Fig.1 denotes a system) work independently, and they get their own results y_i, i=I and II, where x is the observation signal, M the signal model, and θ the model parameter vector. The performance of each system is often very restricted, it is because signal models and signal environments, in general, are complex, such as, some signal model parameters are unknown or the signal is incomplete, or all important parameters of signal model can not be taken into account in a single system when the signal is of a mixture model, or the signal model dose not work with a very low signal to noise ratio, and so on.

In order to improve the signal processing performance, the TISP extracts relevant information from each output of the corresponding system through a "feature extractor" in Fig.1, and the information is exchanged between systems. We call it extrinsic information, since it is a feedback from external systems.

Then, each system works again taking received extrinsic information as new references, generally as a prior condition, and then updates its result.

Thus, the TISP works in such a way of Turbo iteration. As a result, all system's outputs tend to be consistent, and the performance will be continuously improved, just as Turbo Code.

2.2. Iterative Signal Processing Method

From Fig.1, we can see that the TISP is an iterative method. But it is not a self-iteration method as the traditional iterative signal processing method [8] illustrated by Fig.2.

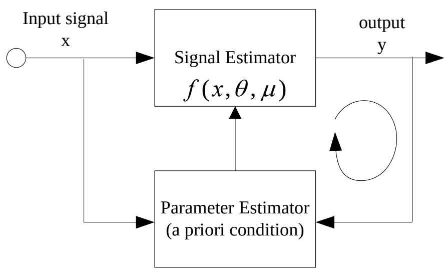


Fig.2. Self-iteration method of signal processing

In a self-iterative algorithm, a set of parameters of signal model or a prior condition of the signal processing system is guessed as the initial step, and then it is updated with last system output. Such a self-iteration with intrinsic feedback information within a system implies a positive feedback, so that the output result falls generally into a local optimum and it is strongly dependent on the initials. For example, the iteration-based non-causal Wiener speech enhancement proposed by Lim and Oppenheim [9], and iterated despeckling for SAR images presented by Oliver [10], and so on, all those are typical self-iterative algorithms.

On the contrary, the Turbo iterative signal processing (Fig.1) is to update a prior condition of a system by the extrinsic information from external system, so that the result tends to seek the global optimum rapidly.

2.3. Multi-Systems Collaborative Method

Another point we can see from Fig.1 is that, Turbo iterative signal processing is a multi-system collaborative working method. However, it is different from the fusion methods for multiple systems.

Fig.3 could describe the decision-level fusion method. In fusion method [11], a proper decision-making is adopted to combine all results of multiple systems in order to improve the performance of decision. Under normal circumstances, there are some limits in each system, so it is difficult to improve the performance in a large extent through a fusion of all these results with defects.

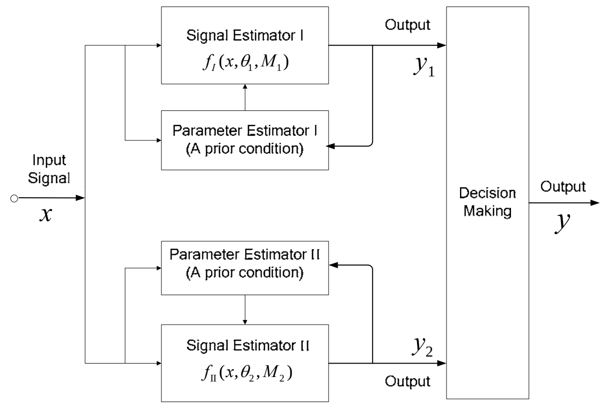


Fig.3. Decision-level fusion method of signal processing

Differently, the Turbo iterative approach (Fig.1) is to allow each system to consult the results of the other systems, and to re-work with the extrinsic references, so that to improve the performance of every system. In comparison with a way of human working in group, decision-level fusion is a "vote" method, while Turbo iterative signal processing is a "discussion" method. The latter, a consensus through discussions, will experientially be much more superior to the former, a judgment through voting.

2.4. Definition of the TISP

To sum up, the method of Turbo-iteration-based signal processing we proposed can be defined as follows.

Turbo iterative signal processing (TISP): In order to improve performance, multiple signal processing systems work in rotation with extrinsic feedback information.

3. TURBO ITERATIVE SAR IMAGE FILTERING

In radar image (such as Synthetic Aperture Radar - SAR image) analysis, it is necessary to reduce multiplicative correlated noise, also called speckle noise. Two global types of speckle analysis have been developed: statistical analysis and wavelet-based space-frequency analysis.

It is much harder to be analyzed and filtered for a multiplicatively noised signal than an additively noised signal. For speckled signals, both statistical filters and wavelet-based filters have some different advantage and also some different intrinsic limitations. That is because each filter imposes some or other signal models (features), and we have no way to take all important signal features into one kind of filter.

3.1. Iterative MAP Estimation

The MAP (maximum a posteriori) estimation takes the probability model of signals into account. From the Bayes' rule the posterior probability density function (PDF) $f_{R|Y}(r|y)$ of the parameter vector r given an observation vector y can be expressed as

$$p_{R|Y}(r|y) = p_{Y|R}(y|r)p_R(r)/p_Y(y).$$

For a given observation vector y, $f_Y(y)$ is a constant, hence the MAP estimate of the parameter vector r is obtained from a maximization of the posterior function:

$$\hat{r}_{MAP} = \arg\max_R p_{R|Y}(r|y) = \arg\max_R \left[p_{Y|R}(y|r) p_R(r) \right].$$

The MAP estimator can benefit from using any prior knowledge about r. For instance, SAR images are assumed as a rather realistic gamma distribution [12]:

$$p_R(r) = \left(\frac{\nu}{\mu}\right)^{\nu} \frac{r^{\nu-1}}{\Gamma(\nu)} \exp\left(-\frac{\nu}{\mu} r\right)$$

where $\Gamma(\cdot)$ denotes Gamma function, μ is the local mean of r in the analyzed zone and ν denotes its degree of heterogeneity:

$$\mu = \frac{1}{N}\sum_{i=0}^{N-1} r_i \quad (1)$$

$$\nu = \frac{\mu^2}{\sigma_r^2}; \quad \sigma_r^2 = \frac{1}{N}\sum_{i=0}^{N-1}(r_i - \mu)^2 \quad (2)$$

A SAR image with multiplicative noise model is a process modeled by $y = r \cdot u$, where y is the observed intensity, r is the underlying reflectivity of ground, u is the unitary-mean speckle contribution uncorrelated with r. Under these hypotheses, the likelihood function of a SAR image can be given by

$$p_{Y|R}(y|r) = \frac{1}{\Gamma(L)}\left(\frac{L}{r}\right)^{L} y^{L-1} \exp\left(-\frac{Ly}{r}\right)$$

where L denotes the equivalent number of looks (ENL) of image data. So, the MAP estimate is given by [10] :

$$\hat{r}_{MAP}(y,\mu,\nu,r) = \arg\underset{r}{zero}\left\{\frac{\nu}{\mu} r^2 + (L+1-\nu) r - Ly\right\} \quad y \neq 0 \quad (3)$$

A MAP filter is to minimize the local Gibbs energy, and the corresponding resolution trends to the local mean μ. In fact, MAP resolution usually stays at a balanced point between μ and y, and the parameter ν indicates the degree of trending to μ:

$$\lim_{\nu\to\infty} r_{MAP} = \mu;$$

$$\lim_{\nu\to 0} r_{MAP} = \frac{L}{L-1} y \approx y, \text{ when } L >> 1$$

Since the parameters of signal model μ and ν are estimated based on speckled signal y within a very limited window N (equations (1) and (2)), an auto-iterative processing in the framework of Fig.2 is applied to improve despeckling performance [10]. Using this nonlinear iterated processing, the MAP algorithm, equations (1), (2) and (3), will be recursive with time k as follows:

$$\cdots \mu^{(k)}, \nu^{(k)} \to r_{MAP}^{(k)} \to \mu^{(k+1)}, \nu^{(k+1)} \to r_{MAP}^{(k+1)} \to \cdots$$

with an initial:

$$r_{MAP}^{(0)} = y$$

As a result, the image sequence $\left\{r_{MAP}^{(k)}\right\}$ tends to be smooth, which means that the noise is reduced while the textures are lost. We take a synthetic speckled intensity image affected by 4-look speckle noise as shown in Fig.4 in out experiments. The results of restoration by iterative MAP estimation are shown in Fig.5. Ultimately, the final result will become a flat image after some hundred iterations [10].

3.2. Iterative Wavelet Shrinkage

Wavelet shrinkage as proposed by Donoho and Johnstone in 1992 [13] has been developed for signals with additive noise, and it has been proved to give good de-noising results. For multiplicative noise a log intensity image is often treaded. The basic idea of wavelet shrinkage is that the large wavelet coefficients usually correspond to textural information, while low coefficients concentrate noise. Therefore a threshold can be applied to wavelet coefficients to distinguish noise from signal.

Within the wavelet decomposition framework, an image data $c(0;i,j)$ at pixel position (i,j) is seen as the sampling at 0 level of the image $f(x,y)$. It is obtained by scalar product (< >) of the image and a scaling function $\phi(x,y)$:

$$c(0;i,j) = \left\langle f(x,y), \phi(x-i, y-j) \right\rangle$$

At scale $a \in (1, A)$, the image approximation is:

$$c(a;i,j) = \left\langle f(x,y), \frac{1}{2^a}\phi\left(\frac{x-i}{2^a}, \frac{y-j}{2^a}\right) \right\rangle$$

The wavelet transform is defined as the difference between two successive approximations:

$$w(a+1;i,j) = c(a;i,j) - c(a+1;i,j)$$

Wavelet shrinkage replaces $w(a;i,j)$ by

$$\hat{w}(a;i,j) = Threshold[w(a;i,j)],$$

where Threshold[w] denotes taking the large wavelet coefficient w, and then we have the result of the wavelet-based shrinkage filter:

$$\hat{c}(0;i,j) \triangleq Shrinkage[c(0;i,j)] = c(A+1;i,j) + \sum_{a=1}^{A} \hat{w}(a;i,j)$$

Since SAR images are modeled by multiplicative noised signal y, the wavelet shrinkage should be applied to the log signal, that is: $c\left(0;i,j\right)\big|_{k=0} = \log y(i,j)$.

Note that the parameter of the threshold in the wavelet shrinkage filter is dependent on *r*. In this case, an auto-iterative processing in the framework of Fig.2 based on wavelet shrinkage using à-trous algorithm can be employed:

$$\log \hat{u}^{(k)} = \log y - \log r_{WT}^{(k)} \tag{4}$$

$$\log \Delta r^{(k)} = Shrinkage\left[\log \hat{u}^{(k)}\right] \tag{5}$$

$$\log r_{WT}^{(k+1)} = \log r_{WT}^{(k)} + \log \Delta r^{(k)} \tag{6}$$

with an initial:

$$r_{WT}^{(0)}(i,j) = average\left[y(i,j)\right]$$

where average[ ] denotes a simple moving average filter.

The above auto-iterative algorithm, equations (4)-(6), converges usually in just *k*=5-10 iterations. Fig.6 shows the results of the synthesis speckled image restoration by iterative wavelet shrinkage, which indicates that the image features are well reserved but the capability of denoising is rather weak, and the estimates are far from the real reflectivity value *r*.

3.3. Turbo Iterative Processing

If we make a fusion in the framework of Fig.3 with the result by the MAP estimation shown as Fig.5 and the result by the wavelet shrinkage shown as Fig.6, the performance for the image restoration becomes better than Fig.5 but worse than Fig.6 at the feature preservation; in the other hand, better than Fig.6 but worse than Fig.5 at noise reduction.

Now, our Turbo iterative processing method in the framework of Fig.1 is employed, where we take:

$$f_1 = \text{MAP Estimation}; \quad f_2 = \text{Wavelet Shrinkage}$$

$$\mathrm{M}_1 = \text{Probability Model}; \quad \mathrm{M}_2 = \text{Spectral Model}$$

$$\theta_1 = \mu, \nu; \quad \theta_2 = r_{WT}(i,j)$$

The Turbo iterative despeckling algorithm can be described as

$$k = 1, 2, \cdots$$

$$\mu^{(k)} = \frac{1}{N}\sum_{i=0}^{N-1} \hat{r}_{WT}^{(k-1)}; \quad \nu^{(k)} = \frac{(\mu^{(k)})^2}{\sigma^2_{\hat{r}_{WT}^{(k-1)}}} \tag{7}$$

$$\hat{r}_{MAP}^{(k)}\left[y, \mu^{(k)}, \nu^{(k)}, \hat{r}_{WT}^{(k-1)}\right]$$

$$\log u^{(k)} = \log y - \log \hat{r}_{MAP}^{(k)} \tag{8}$$

$$\log \Delta r^{(k)} = Shrinkage\left[\hat{u}^{(k)}\right]$$

$$\log \hat{r}_{WT}^{(k)} = \log \hat{r}_{MAP}^{(k)} + \log \Delta r^{(k)} \tag{9}$$

This TISP algorithm propagates the information of FilterⅡ (wavelet shrinkage) to Filter Ⅰ (MAP estimator) as a part of a priori knowledge in MAP algorithm through (7), and in the other way, it propagates the information of Filter Ⅰ to FilterⅡ as the reference image in shrinking algorithm through (8) and (9). This information exchange is of prime importance for the efficiency of TISP denoising. Our experimental results have shown that, in fact, the auto-iterative MAP algorithm tends to provide a too smooth image $\hat{r}_{MAP}^{(k)}$ while losing so much texture in the image *y* (Fig.5); on the contrary, the auto-iterative wavelet shrinkage tends to preserve texture but to leave a lot of noise in the estimated image $\hat{r}_{WT}^{(k)}$ (Fig.6). The Turbo iterative algorithm balances well these two trends (Fig.7).

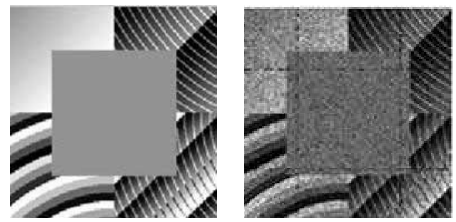
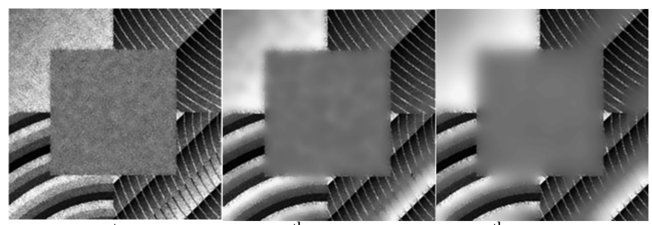
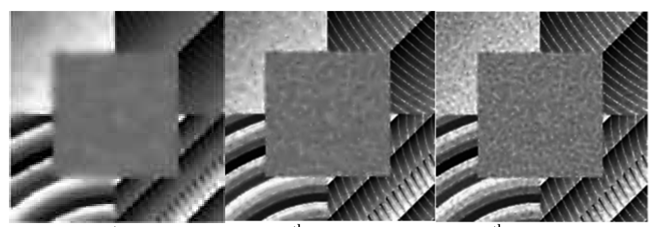


Fig.4 (a) Synthesis image; (b) Synthesis speckled image



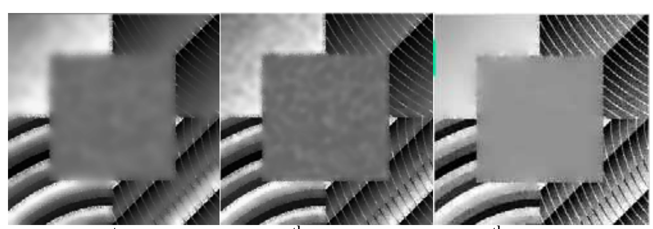
(a) 1st iteration; (b) 5th iteration; (c) 20th iteration

Fig.5 Restoration of Iterative MAP Estimation



(a) 1st iteration; (b) 5th iteration; (c) 20th iteration

Fig.6 Restoration of Iterative Wavelet Shrinkage Filter



(a) 1st iteration; (b) 5th iteration; (c) 10th iteration

Fig.7 Turbo Iterative Restoration

4. TURBO ITERATIVE SPEECH ENHANCEMENT

Now, we consider speech enhancement in the framework of the TISP.

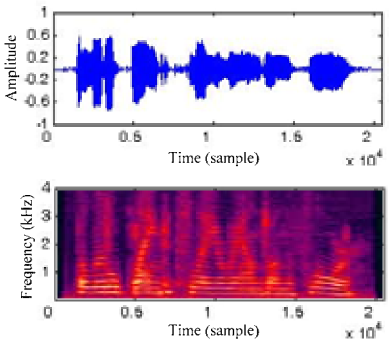


Fig.8 (a) Clean utterance in time and frequency domains;

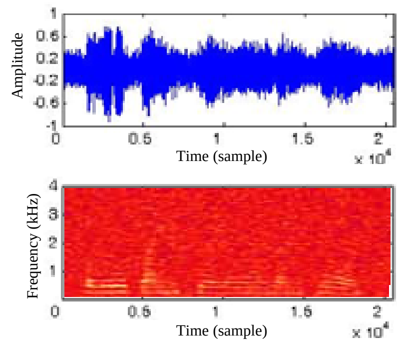


Fig.8 (b) noisy utterance in time and frequency domains

Speech enhancement approaches are often used in the complex environments, such as high levels of ambient noise, or lack of model parameters, etc. We have a variety of different speech enhancement methods based on different speech models developed over the past several decades, which have their own advantages and different limitations. Under the same principle of the Turbo iterative image restoration presented in the last section, we can use a Kalman filter with the voice generation model and the wavelet threshold filter with the short-term spectrum model for speech enhancement. Both filters work in rotation and each one takes some feedback information from the other filter as a priori condition. In the TISP framework of Fig.1, we have

$$f_1 = \text{Kalman Filter}; \qquad f_2 = \text{Wavelet Shrinkage}$$
$$\mathrm{M}_1 = \text{Voic Generation Model}; \qquad \mathrm{M}_2 = \text{Spectral Model}.$$

The Turbo iterative speech enhancement can be described as:

$$\begin{aligned} \hat{s}_{KF}^{(k)} &= KalmanFilter\left[\hat{s}_{KF}^{(k-1)} + \Delta\hat{s}_{WT}^{(k-1)}\right] \\ \Delta\hat{s}_{WT}^{(k)} &= WaveletShrinkage\left[y - \hat{s}_{KF}^{(k)}\right] \end{aligned} \qquad (10)$$

where s denotes clean speech signal, y denotes observation of noised speech signal;

$$x_{KF}^{(k)} = KalmanFilter\left[x_{KF}^{(k-1)}\right]$$

denotes the output $x_{KF}^{(k)}$ of the Kalman filter at k with the output $x_{KF}^{(k-1)}$ at $k-1$; and

$$y = WaveletShrinkage\left[x\right]$$

denotes the output y of the wavelet shrinkage filter with the input x.

Our experiments take 10 speech utterances with three different speakers and 30 sec of speech. The speech is sampled at 8 kHz and quantized to 16 bits. And Computer-generated stationary white Gaussian noise is artificially added at 0 dB SNR (Fig.8), its variance being assumed to be perfectly known. A frame size of 32 ms with 50% overlap is used.

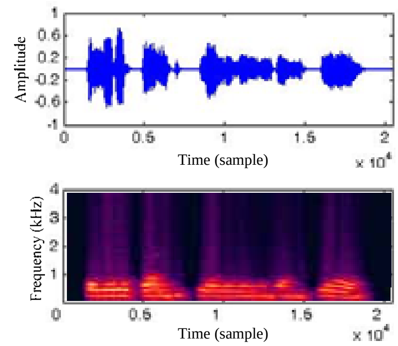


Fig.9 Speech enhancement by auto-iterative Kalman filter

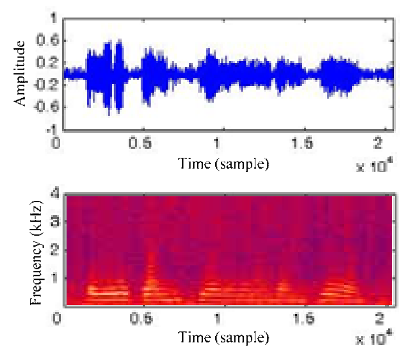


Fig.10 Speech enhancement by auto-iterative wavelet shrinkage

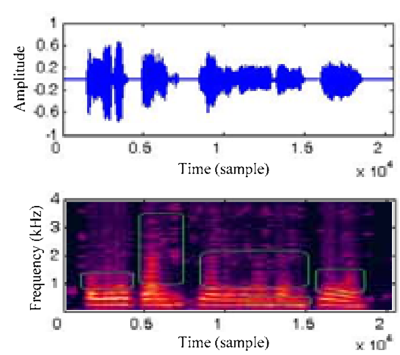


Fig.11 Speech enhancement by TISP method

Fig.9 shows the result of enhancement of the noisy utterance in Fig8b by using auto-iterative Kalman filter and Fig.10 shows the result by using auto-iterative wavelet shrinkage filter. Fig.11 shows the result of enhancement by using the TISP method (10).

We can see from the experiment results that the auto-iterative Kalman filter tends to provide a too suppressed speech signal while some unvoiced speech signal is filtered as noise (Fig.9). On the contrary, the auto-iterative wavelet threshold tends to extract detail information but a lot of noise rest in the enhanced speech signal (Fig.10). The Turbo iterative algorithm balances well these two trends. From the example spectrum in Fig.11, the TISP method can effectively suppress the noise, and it also helps preserve weak speech segment information (look at the highlighted rectangular areas in Fig.11).

5. TURBO ITERATIVE IMAGE PROCESSING

We have applied the Turbo iterative method to image processing, including SAR image segmentation, classification, pixel-level image fusion and so on.

For SAR image segmentation, we used a Bayesian estimator with a Markov Random Field model and the Iterative Conditional Estimation algorithm with a Multi-Level Logistic model. Both systems work in rotation and each one takes some feedback information from the other system as a priori condition. Using this TISP framework of Fig.1, we got a high quality of SAR image segmentation result [4].

For SAR pixel-level image fusion, we use a wavelet-based fusion algorithm with the spectrum model and a model-based fusion method with the probability model. We make them work in the Turbo iterative way. The fusion results for SAR images show a good performance [5].

6. SUMMARY AND DISCUSSION

Turbo iterative signal processing (TISP) we proposed is a multi-systems collaborative working method: each system works in rotation with the extrinsic feedback information extracted from the other systems. For complex signal processing problems, the TISP utilizes several simple algorithms to work collaboratively in the way of Turbo iteration, instead of a unique complex algorithm with a mixture signal model.

We have presented several applications of the TISP on image processing and speech processing:

SAR image filtering by the TISP: system 1 (ref. Fig.1) is a MAP estimator using the probability model of SAR images with the parameters, local mean and local variance, extracted from the output of system 2; system 2 is a wavelet-based shrinking filter using the spectral model with the parameter, the speckle noise space, extracted from the output of system 1.

Speech enhancement by the TISP: system 1 is a Kalman filter using a dynamic predictive model of a speech signal with the parameters, the coefficients of the state equation of the speech signal, extracted from the output of system 2; system 2 is a wavelet-based filter using the short-time spectral model with the parameter, the noise space, extracted from the output of system 1.

SAR image processing, considering image segmentation and fusion, by the TISP: two different signal models are used in two corresponding algorithms, and the two systems work in Turbo iteration.

All application examples above show a very good performance after 5-10 iterations.

In fact, the TISP presents an effective machine-learning method: making a discussion between several independent systems through Turbo iteration.

This work is done during H. Sun's sabbatical period at Telecom-ParisTech.

REFERENCES

[1] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon-limit Error-correcting Coding and Decoding: Turbo Codes," *Proc.ICC'93*, Geneva, pp.1064-1070, 1993.

[2] H. Sun, and H. Maître, "L'analyse radiometrique multiresolution et son application au filtrage du chatoiement," Depôt légal: 2000. 4ème, ENST–Paris, ISSN0751-1345, Nov. 2000

[3] H. Sun, H. Maître, and B. Guan, "Turbo Image Restoration," *Proceedings of Seventh International Symposium on Signal Processing and Its Applications*, Paris, Vol.1, pp.417-420, July 1-4, 2003.

[4] Y.F. Cao, H. Sun, and X. Xu, "An Unsupervised Segmentation Method Based on MPM for SAR Images," *IEEE Geoscience and Remote Sensing Letters*, Vol.2, No.1, pp.55-58, January 2005.

[5] C. He, G.S. Xia, W.M. Peng, and H. Sun, "SAR Image Fusion Based on Turbo Iterative," The 4th International Symposium on Multi-spectral Image Processing and Pattern Recognition, Wuhan, Vol.6043, *SPIE*, October, 2005.

[6] B. Guan, H. Sun, "Turbo Iterative Estimation of Singularity Structure in SAR Image based on Wavelet-domain Hidden Markov Models", ICASSP2005, USA, Vol.2, pp.601-604, March 2005

[7] H. Dong, and H. Sun, "Turbo Iterative Speech Enhancement", *ICASSP2008*, submitted, 2008

[8] S.V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*, John Wiley & Sons, Ltd., 2000.

[9] J.S. Lim, and A.V. Oppenheim, "All-pole modelling of degraded speech," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.26, No.3, pp.197-210, 1978.

[10] C. Oliver, and S. Quegan, *Understanding Synthetic Aperture Radar Images*, Artech House Inc., 1998.

[11] I. Bloch, *Fusion d'informations en traitement du signal et des images*, LAVOISIER, 2005

[12] E. Nezry, A. Lopès, C.N. Ducrot-Gambart, J.S. Lee, "Supervised Classification of K-distributed SAR images of natural targets and probability of error estimation," IEEE trans. on Geoscience and Remote Sensing, 34(5), pp.1233-1242, 1996.

[13] D.L. Donoho, and I.M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," Biometrika.J., vol.81, pp.425--455, 1994.