A Fast Learning Algorithm for Blind Data Fusion Using a Novel L_2-Norm Estimation

Youshen Xia and Henry Leung, *Member, IEEE*

***Abstract*—In this paper, a novel L_2-norm estimation method for blind data fusion under noisy environments is proposed and a fast learning algorithm is developed to implement the proposed estimation method. The proposed learning algorithm is proved to be globally exponentially convergent to an optimal fusion weight vector. In addition, the proposed learning algorithm has lower computation complexity than the existing cooperative learning algorithm based a L_1-norm estimation method. Compared with other estimation methods, the proposed estimation method can be effectively used in the blind image fusion. Application examples of image fusion show that the proposed learning algorithm is able to fast obtain more accurate solutions than several conventional algorithms.**

***Index Terms*—Data fusion, blind image fusion, novel L_2 estimation, fast algorithm.**

I. Introduction

With the availability of multiple sensory devices, multisensor fusion has received considerable attention in recent years. The main advantage of data fusion is that by integrating complementary information from different sources, the fused performance can be improved [1]. As a result, the multisensor data fusion has found widespread applications in many fields such as machine vision, signal and image processing, nonlinear modeling and identification, etc. [2]–[5]. A fusion process can be carried out at various levels such as signal; image; and feature. Signal level fusion combines raw data received from different sensors to provide a consensus signal. Image level fusion may be viewed as a special case of the signal level fusion with application to images [27].

Various statistically optimal signal/pixel level fusion techniques have been developed, such as the minimum variance estimation [6], the maximum likelihood technique [7], the wavelet transform technique [8], [9], the linearly constrained least squares (LCLS) method [10], [11], and the covariance intersection (CI) method [28], [29]. Under the Gaussian assumption, the maximum likelihood estimate is equivalent to the minimum variance method given the sensor scaling vector and noise covariance. They are found effective when the covariance of the sensory information is given. The LCLS method and the wavelet transform method can overcome the unavailable covariance information under Gaussian noise environment. However, the Gaussian assumption is for the mathematical convenience and it does not generally hold in practical applications. Moreover, these methods only deal with the cases of known scaling coefficients. To handle non-Gaussian noise problems, a linearly constrained L_1 estimation method was proposed in the paper [12]. Although the L_1 estimation-based fusion method can deal with non-Gaussian noise problems, it is suitable for dealing with the cases of known scaling coefficients. In practice, the scaling coefficients are unknown. One example is the blind signal fusion problem [2], [5], [7], another example is the image fusion problem [13], [14]. Recently, an L_1 estimation-based fusion method was presented by using a relaxed non-convex constraint technique [15]. This method can obtain an approximate unbiased fusion estimate. Because it has to minimize a non-smooth cost function, the resulting cooperative learning algorithm has very high computational complexity and thus has a slow speed.

Based on a novel L_2-norm estimation method, this paper proposes a fast learning algorithm for blind fusion of multiple sensor data under noisy environments. The proposed learning algorithm is shown to be globally exponentially convergent to an optimal fusion weight vector. The proposed learning algorithm minimizes a convex quadratic cost function of the linearly fused information and thus has significantly lower computational complexity than the existing cooperative learning algorithm based on a constrained L_1-estimation method. Compared with other statistical data fusion methods, the proposed method can be effectively used in blind image fusion problems without an Gaussian noise assumption. Application examples show that the proposed learning algorithm can fast obtain more accurate solutions than other statistical fusion algorithms.

II. Blind Fusion Problem and Estimation

Consider a multi-sensor system with K ($K \geq 2$) sensors. Let the kth sensor measurement be expressed as

$$x_k(t) = a_k s(t) + n_k(t), (k = 1, ..., K; t = 1, ..., N), \quad (1)$$

where a_k is an unknown sensor coefficient, N is the number of sensor measurements, K ($K \geq 2$) is the sensor number, $s(t)$ is the original signal with $\gamma_1 \leq s(t) \leq \gamma_2$, and $x_k(t)$

Manuscript received August 28, 2013; accepted September 14, 2013. Date of publication September 9, 2013; date of current version January 7, 2014. This work was supported by the National Natural Science Foundation of China under Grant 61179037 and 60875085, and the Doctoral Project of the Ministry of Education of China (20133514110010). The associate editor coordinating the review of this paper and approving it for publication was Dr. Amitava Chatterjee.

Y. Xia is with the College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China (e-mail: ysxia@fzu.edu.cn).

H. Leung is with the Department of Electrical and Computer Engineering, University of Calgary, Calgary T2N 1N4, Canada (e-mail: leungh@ucalgary.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSEN.2013.2282693

and $n_k(t)$ represents the received image signal and noise at the kth sensor, respectively. (1) is a popular blind fusion model because it is widely discussed in literatures [2], [5], [13]–[15]. Furthermore, the sensor sources may be signals where t indicates a time series and N is the number of sensor measurements within the time window, and sensor sources may be images where t indicates the spatial coordinates of a pixel and N denotes the number of sensor measurements within the pixel coordinates window. Using matrix and vector notations, (1) can be written as

$$\mathbf{x}(t) = \mathbf{a}s(t) + \mathbf{n}(t), \tag{2}$$

where $\mathbf{a} = [a_1, ..., a_K]^T$, $\mathbf{x}(t) = [x_1(t), ..., x_K(t)]^T$, and $\mathbf{n}(t) = [n_1(t), ..., n_K(t)]^T$. Let the fusion scheme be a linear combination of sensor measurements given by

$$u(t) = \sum_{k=1}^{K} w_k x_k(t) = \mathbf{w}^T \mathbf{x}(t), \tag{3}$$

where $\mathbf{w} = [w_1, w_2, ..., w_K]^T$ is the weighting vector. The objective of blind fusion is to estimate $s(t)$ by combining the sensory information in the case of unknown coefficients $\{a_k\}$.

Under an assumption of Gaussian noise, the minimum variance estimation method finds an optimal fusion weight $\mathbf{w}$ such that

$$\begin{aligned} &\text{minimize } \sum_{t=1}^{N} (\mathbf{w}^T \mathbf{x}(t) - s(t))^2 \\ &\text{subject to } \mathbf{a}^T \mathbf{w} = 1. \end{aligned} \tag{4}$$

Since the covariance information is usually not available, the minimum variance method can be modified by the linearly constrained least squares (LCLS) fusion method which finds an optimal fusion weight $\mathbf{w}$ satisfying

$$\begin{aligned} &\text{minimize } f_1(\mathbf{w}, \mathbf{a}) = \mathbf{w}^T R \mathbf{w} \\ &\text{subject to } \mathbf{a}^T \mathbf{w} = 1, \end{aligned} \tag{5}$$

whose optimal solution is given by

$$\mathbf{w}^*_{LS} = R^{-1}\mathbf{a}/(\mathbf{a}^T R^{-1} \mathbf{a}), \tag{6}$$

where $R = \sum_{t=1}^{N} \mathbf{x}(t)^T \mathbf{x}(t)/N$. Since R is quite often ill conditioned and close to singular when the dependency exists between the noise sources, a poor solution may be obtained by (6). To overcome this shortcoming of numerical instability, a neural LCLS fusion algorithm was developed. However, the neural LCLS fusion algorithm assumes known scaling coefficients. Thus both the minimum variance fusion algorithm and the LCLS fusion algorithm suffer from the local minima problem in finding the optimal fusion due to the non-convex constraint $\mathbf{a}^T \mathbf{w} = 1$. Also, the wavelet-based fusion algorithm is popular but it may loose much of contrast information in blind fusion and also require an assumption of Gaussian noise.

Recently, a constrained L_1 fusion estimation method [12] was developed to handle the non-Gaussian noise problem by minimizing

$$\begin{aligned} &\text{minimize } \sum_{t=1}^{N} |\mathbf{w}^T \mathbf{x}(t) - s(t)| \\ &\text{subject to } \mathbf{a}^T \mathbf{w} = 1, \quad \gamma_1 \le s(t) \le \gamma_2 \end{aligned} \tag{7}$$

Furthermore, to overcome difficulty of the non-convex difficulty in blind fusion, a relax constraints-based L_1 fusion estimation method [15] is presented by solving

$$\begin{aligned} &\text{minimize} \quad f_2(\mathbf{w}, \mathbf{a}, \mathbf{z}) = \|X\mathbf{w} - X\mathbf{a} - \mathbf{z}\|_1 \\ &\text{subject to} \quad \mathbf{z} \in \mathbf{\Omega}_1, \end{aligned} \tag{8}$$

where $\mathbf{w}$ is called the weight regression vector, $\mathbf{z}$ is called the residual regression vector, $\|\cdot\|_1$ denotes l_1 norm, $X = [\mathbf{x}(1), ..., \mathbf{x}(N)]^T$, residual error set $\mathbf{\Omega}_1 = \{\mathbf{z} \in R^N \mid \mathbf{l}^1 \le \mathbf{z} \le \mathbf{h}^1\}$. The optimal weight vector is given by

$$\hat{\mathbf{w}}^*_{LAD} = \mathbf{w}^*/(\mathbf{w}^*)^T \mathbf{a}^*$$

where $(\mathbf{w}^*, \mathbf{a}^*, \mathbf{z}^*)$ is an optimal solution of (8). To solve the non-smooth optimization problem (8), a cooperative learning algorithm was developed in [15]. However, it has a slow speed due to a high computational complexity.

III. Novel Optimal Fusion Estimation

In this section we introduce a fast learning algorithm for blind fusion based on a novel L_2 estimation approach.

The proposed novel L_2 estimation approach is to find an optimal fusion weight vector $\mathbf{w}^*$ satisfying

$$\begin{aligned} &\text{minimize} \frac{1}{2} \sum_{t=1}^{N} (\mathbf{w}^T \mathbf{x}(t) - s(t))^2 \\ &\text{subject to} \quad \mathbf{a}^T \mathbf{w} = 1, \quad \gamma_1 \le s(t) \le \gamma_2 \end{aligned} \tag{9}$$

To remove the non-convex constraint $\mathbf{a}^T \mathbf{w} = 1$, we first analyze the relationship between the fusion weight and the scaling coefficient. On one side, we have $E[\mathbf{x}(t)] = \mathbf{a}E[s(t)]$ and thus

$$E[s(t)] = \frac{\mathbf{a}^T E[\mathbf{x}(t)]}{\|\mathbf{a}\|_2^2}$$

where $\|\cdot\|_2$ denotes l_2 norm. On the other side, an unbiased fusion estimate implies that $E[s(t)] = \mathbf{w}^T E[\mathbf{x}(t)]$. Then

$$\frac{\mathbf{a}^T E[\mathbf{x}(t)]}{\|\mathbf{a}\|_2^2} = \mathbf{w}^T E[\mathbf{x}(t)]. \tag{10}$$

(10) implies that there is a linear relationship between the fusion weight and the scaling coefficient. That is,

$$\mathbf{w} = \rho \frac{\mathbf{a}}{\|\mathbf{a}\|_2^2}.$$

Note that $\mathbf{a}^T \mathbf{w} = 1$. Then $\rho = 1$ and thus $\mathbf{w} = \mathbf{a}/\|\mathbf{a}\|_2^2$.

Based on the analysis above, we reformulate (9) as a quadratic convex regression problem:

$$\begin{aligned} &\text{minimize} \quad f_3(\mathbf{w}, \mathbf{z}) = \frac{1}{2}(\|X\mathbf{w} - \mathbf{z}\|_2^2 + \eta \|\mathbf{w}\|_2^2) \\ &\text{subject to} \quad \mathbf{z} \in \mathbf{\Omega}, \end{aligned} \tag{11}$$

where $(\mathbf{w}, \mathbf{z})$ is the regression vector, X is defined in (8), $\mathbf{\Omega} = \{\mathbf{z} \in R^N \mid \gamma_1 \mathbf{e} \le \mathbf{z} \le \gamma_2 \mathbf{e}\}$, and $\eta > 0$ is a regularization parameter. Now, let $(\mathbf{w}^*, \mathbf{z}^*)$ be the optimal solution of (11) and let

$$\hat{a}^*_k = E[x_k(t)]/(\gamma \sum_{i=1}^{N} \mathbf{z}^*_i/N), \ k = 1, ..., K \tag{12}$$

where γ is a design constant. Then the new optimal fusion weight is defined as

$$\mathbf{w}^*_{new} = \hat{\mathbf{a}}^*/\|\hat{\mathbf{a}}^*\|_2^2 = \mathbf{w}^*/(\mathbf{w}^*)^T\hat{\mathbf{a}}^* \tag{13}$$

where $\hat{\mathbf{a}}^* = [\hat{a}_1^*, \hat{a}_2^*, ..., \hat{a}_K^*]^T$, and the optimal fusion estimate is given by

$$u^*(t) = (\mathbf{w}^*_{new})^T\mathbf{x}(t). \tag{14}$$

We now show that the proposed optimal fusion is unbiased. Because the additional noise is white,

$$a_k = E[x_k(t)]/E[s(t)],\ k = 1, ..., K.$$

Note that the cost function defined in (11) may approximate to the cost function defined in (9) when parameter η is small enough. Then $\mathbf{z}^* \approx [s(1), ..., s(N)]$. Thus the arithmetic average of $\mathbf{z}^*$ approximates the mean of $s(t)$. That is, $E[s(t)] \approx \frac{1}{N}\sum_{k=1}^N z_k^*$. Therefore, there exists a constant γ such that $E[s(t)] = \gamma\sum_{k=1}^N z_k^*/N$. It follows that

$$\hat{a}_k^* = E[x_k(t)]/(\gamma\sum_{i=1}^N \mathbf{z}_i^*/N) = E[x_k(t)]/E[s(t)] = a_k \quad (k = 1, ..., K)$$

Then

$$E[u^*(t)] = E[(\mathbf{w}^*_{new})^T\mathbf{x}(t)] = (\hat{\mathbf{w}}^*_{new})^T\hat{\mathbf{a}}^*E[s(t)] = E[s(t)]$$

since

$$(\mathbf{w}^*_{new})^T\hat{\mathbf{a}}^* \quad = \frac{(\hat{\mathbf{a}}^*)^T}{\|\hat{\mathbf{a}}^*\|_2^2}\hat{\mathbf{a}}^* = 1.$$

Therefore, the proposed optimal fusion is unbiased if γ is properly chosen.

IV. Fast Learning Algorithm

Many algorithms were developed for solving the constrained least-square problems [16]–[22]. Among them, the CVX algorithm is viewed as a popular numerical algorithm since it has both low computational complexity and fast convergence. When solving (11), however, the CVX algorithm needs at least $O((N+K)^2)$ multiplication operation per iteration. Based on the special structure of (14), we introduce a fast learning algorithm with $O(N+K)$ multiplication operation per iteration only.

By the optimality condition [16] we see that $(\mathbf{w}^*, \mathbf{z}^*)$ is an optimal solution of (11) if and only if

$$(\mathbf{w}-\mathbf{w}^*)^T\frac{\partial f_3(\mathbf{w}^*,\mathbf{z}^*)}{\partial\mathbf{w}} + (\mathbf{z}-\mathbf{z}^*)^T\frac{\partial f_3(\mathbf{w}^*,\mathbf{z}^*)}{\partial\mathbf{z}} \geq 0 \quad \forall\mathbf{w}\in R^K,\ \forall\mathbf{z}\in\Omega \tag{15}$$

Furthermore, by the projection theorem [15] we know that $(\mathbf{w}^*, \mathbf{z}^*)$ satisfies

$$\begin{cases} \frac{\partial f_3(\mathbf{w},\mathbf{z})}{\partial\mathbf{w}} = 0 \\ g(\mathbf{z} - \frac{\partial f_3(\mathbf{w},\mathbf{z})}{\partial\mathbf{z}}) = \mathbf{z} \end{cases} \tag{16}$$

where $g(\mathbf{z})$ is a projection function from R^N to the set Ω, given by $g(\mathbf{z}) = [g(z_1), ..., g(z_N)]^T$ and for $i = 1, ..., N$

$$g(\mathbf{z}_i) = \begin{cases} \gamma_1 & z_i < \gamma_1 \\ z_i & \gamma_1 \leq z_i \leq \gamma_2 \\ \gamma_2 & z_i > \gamma_2 \end{cases}$$

Note that

$$\frac{\partial f_3}{\partial\mathbf{w}} = (X^TX+\eta I)\mathbf{w} - X^T\mathbf{z},\ \frac{\partial f_3}{\partial\mathbf{z}} = -(X\mathbf{w}-\mathbf{z})$$

Then $(\mathbf{w}^*, \mathbf{z}^*)$ is an optimal solution of (11) if and only if $(\mathbf{w}^*, \mathbf{z}^*)$ satisfies the following system of two equations

$$\begin{cases} (X^TX+\eta I)\mathbf{w} = X^T\mathbf{z} \\ g(X\mathbf{w}) = \mathbf{z} \end{cases} \tag{17}$$

Then $(X^TX+\eta I)\mathbf{w} = X^Tg(X\mathbf{w})$. We thus present a novel learning algorithm for solving (11) as follow

State equation

$$\frac{d\mathbf{w}(t)}{dt} = \lambda\{-(X^TX+\eta I)\mathbf{w}(t) + X^Tg(X\mathbf{w}(t))\} \tag{18a}$$

where $\lambda > 0$ is a design constant.

Output equation

$$\mathbf{z}(t) = g(X\mathbf{w}(t)) \tag{18b}$$

Based on (18), we propose the following neural fusion algorithm for blind fusion of multiple sensor data as follows:

Given parameters γ, η, and λ.

Step 1. Solve (11) by the proposed learning algorithm (18) and let $(\mathbf{w}^*, \mathbf{z}^*)$ be output in steady state.

Step 2. Compute fusion weight vector

$$\mathbf{w}_{new} = \mathbf{w}^*/(\mathbf{w}^*)^T\hat{\mathbf{a}}^* \tag{19}$$

where

$$\hat{\mathbf{a}}^* = E[\mathbf{x}(t)]/(\gamma\sum_{i=1}^N\mathbf{z}_i^*/N)$$

Step 3. Compute fusion estimate:

$$u_{new}(t) = (\mathbf{w}_{new})^T\mathbf{x}(t). \tag{20}$$

Remark 1: In the proposed algorithm, there are three parameters to be designed. For an unbiased fusion estimate, we should take a proper design parameter γ. Note that

$$E[s(t)]\frac{1}{K}\sum_{i=1}^K a_i = \sum_{i=1}^K E[x_i(t)]$$

and

$$\sum_{i=1}^N \mathbf{z}_i^*/N \approx \sum_{i=1}^K E[x_i(t)].$$

Then it is a good choice if we take the design parameter as

$$\gamma = 1/(\frac{1}{K}\sum_{i=1}^K a_i).$$

Also, from the proof of Theorem 1 we see that $\eta\lambda$ is a convergence ratio of the proposed learning algorithm. On the other side, the regularization parameter η has to be small enough for an optimal weight. Therefore, we take design parameters η and λ satisfying $\lambda \geq 1/\eta$ with a small η.

As for the convergence of the proposed learning algorithm, we have the following result.

Theorem 1. The proposed learning algorithm is globally exponentially convergent to the optimal fusion weight vector.

Proof. The proof is given in the Appendix.

Finally, we analyze the computational complexity of the proposed learning algorithm. To do that, we first define the algorithm complexity as the total number of multiplications/divisions per iteration. Furthermore, we employ the total asymptotic complexity in terms of the asymptotically tight bound function $\theta(\cdot)$ [26], defined as

$$\begin{gathered}\theta(g(n)) = \{f(n)| \\ \text{there exist positive constants } c_1, c_2, \text{and } n_0 \\ such\ that\ 0 \leq c_1 g(n) \leq f(n) \leq c_2 g(n)\ \forall\ n \geq n_0.\}.\end{gathered}$$

Note that the proposed learning algorithm finds the optimal solution by tracking the solution trajectory of (18) based on its discrete-time form:

$$\mathbf{w}(t+1) = \mathbf{w}(t) + h_t[C\mathbf{w}(t) + X^T g(X\mathbf{w}(t))]$$

where $C = -(X^T X + \eta I)$ and h_t is a step length factor. So, we need K^2 and $2KN$ multiplication operations for computing $C\mathbf{w}$ and $X^T g(X\mathbf{w})$ per iteration, respectively. Thus the proposed learning algorithm has the total number of multiplications being $K^2 + 2KN + K$ and hence its asymptotic complexity is $O(N+K)$. By contrast, in a similar analysis we know that the cooperative learning algorithm proposed in [15] has the total numbers of multiplications $KN^2 + (8+4K)N$ per iteration and its asymptotic complexity is $\theta(KN^2)$. So, the proposed algorithm has a lower computational complexity than the CLAD fusion algorithm. Furthermore, it can be seen that the CVX algorithm given in [16] needs $O((N+K)^2)$ multiplication operation per iteration. Thus, the proposed learning algorithm also has a lower computational complexity than the CVX algorithm. We will show in computed examples that the proposed learning algorithm has a faster speed than the CVX algorithm. This is more significant in real-time image fusion. For example, if an image is 256×256 pixels, then $N = 65536$. Both the CLAD fusion algorithm and the CVX algorithm require at least 4.3×10^9 multiplication operation per iteration. By contrast, the proposed learning algorithm then requires 6.6×10^4 multiplication operation per iteration. Furthermore, we will show in application examples that the proposed learning algorithm has a faster speed than the CVX algorithm.

V. Application to Image Fusion

In this section, we apply the proposed learning algorithm to image fusion. Image fusion can be used to enhance the quality of an image so that more reliable segmentation can be taken place and more discriminating features can be extracted for image processing. In the experiments, synthetic and real images are adopted as test images. Among them, the synthetic images are used to verify whether the proposed fusion algorithm is consistent with its performance, and the real images are used to examine how well the proposed algorithm works. The experiments are conducted in MATLAB.

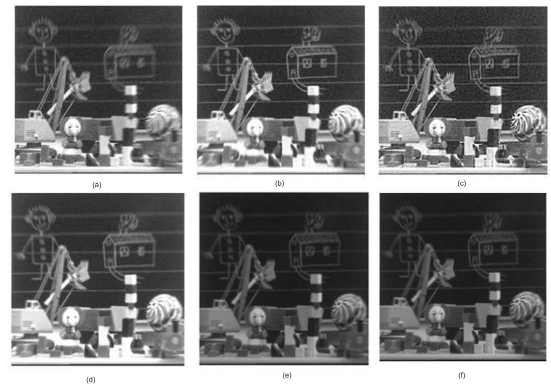

Fig. 1. Image fusion of example 1: (a) and (b) Noisy blurred images; (c) the fused image by the wavelet algorithm; (d) the fused image by the LCLS algorithm; (e) the fused image by the LAD algorithm; (f) the fused image by the proposed fusion algorithm.

Because an image pixel value usually satisfies $0 \leq s(t) \leq L$ with $s(t) \not\equiv 0$, in the proposed learning algorithm we set parameters $\gamma_2 = L$ and $\gamma_1 = \epsilon_t$ where $\epsilon_t = min\{s(t)\}$ for $s(t) > 0$ and $\epsilon_t = 0$ for $s(t) = 0$. For example, we take γ_1 as a non-negative random number. When performing the proposed learning algorithm for the fusion weight, we employ the numerical ODE23 software which is the discrete-time version of the continuous-time learning algorithm defined in (18) where parameters $\eta = 10^{-8}$ and $\lambda = 150$ are taken.

In the previous two examples, we study synthetic images as test images. To illustrate the better performance of the proposed learning fusion algorithm, we compare it with the L_1 estimation-based cooperative learning fusion (LAD) algorithm, the LCLS fusion algorithm, and the wavelet-based fusion algorithm. The performance of the algorithm is evaluated by visual quality, the image quality index Q [23], and the weighted quality index QW [24]. The image quality index is designed to measure image distortion in three factors: loss of correlation, luminance distortion, and contrast distortion. The weighted quality index is designed to measure image distortion by using a convex combination of the image quality index.

Example 1. Consider the problem of multi-focus image fusion. Three registered images with different focal points are taken from the image fusion research of the Lehigh University. They are 8-bit grey level images with 512 by 512 pixels. Each image has a different scaling coefficient

$$a_i = 1 + 0.2i \ \ (i = 1, 2, 3), \tag{21}$$

and the noise is a white uniformly distributed in area [0, 1] with a variance of $\sigma^2 = 50$. Fig. 1 (a) and (b) display the noisy blur images. Fig. 1 (c) and (d) display fused images obtained by the wavelet-based fusion algorithm and the LCLS fusion algorithm, respectively. It can be seen that both the fused images have lost the contrast information. We now perform the L_1 estimation-based learning algorithm and the proposed learning algorithm with parameters $L = 255$, $N = 512$, $\alpha = 4.2$, and $\gamma = 2$. Fig. 1 (e) and (f) display the fused images

TABLE I
PERFORMANCE COMPARISON OF FOUR ALGORITHMS IN EXAMPLE 1

Method	wavelet	LCLS	LAD	Proposed algorithm
CPU time	0.9850	0.8340	44.6570	0.187
Q	0.6926	0.6060	0.8310	0.8286
QW	0.7626	0.5830	0.5530	0.5542

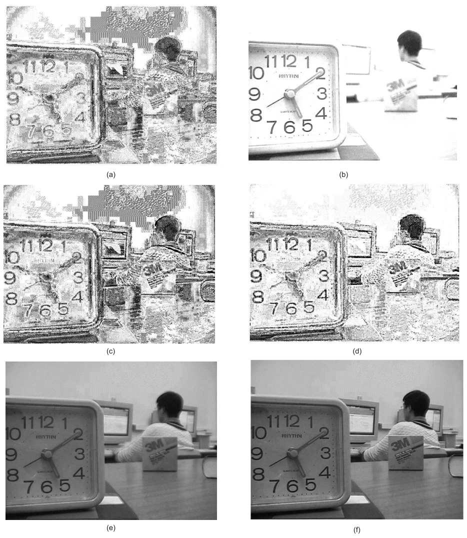

Fig. 2. Blind fusion of example 2: (a) and (b) Noisy blurred images; (c) the fused image by the wavelet algorithm; (d) the fused image by the LCLS algorithm; (e) the fused image by the L_1 estimation-based cooperative learning algorithm; (f) the fused image by the proposed learning algorithm.

obtained by the L_1 estimation-based cooperative learning algorithm and the proposed learning algorithm, respectively. As seen, the loss of contrast information is greatly reduced by the two algorithms. Furthermore, Table I lists the image metric and computation-time results of the four algorithms where CPU time is denoted by second. From Table I we see that the proposed learning algorithm obtains the best result on image quality Q. Moreover, the proposed learning algorithm is much faster than other three fusion algorithms. Finally, we use the CVX algorithm for the optimal fusion weight defined in (14). The CVX algorithm can give better result similar to the proposed learning algorithm but takes computation time being 1.53 second. Thus, the proposed learning algorithm also has a faster speed than the CVX algorithm.

Example 2. Consider four digital camera images with different focal points taken by using Sony digital camera from the image fusion research of the Lehigh University. They are 8-bit grey level images with 480 by 640 pixels. Each image has a different scaling coefficient

$$a_i = 1 + 0.5i \ \ (i = 1, 2, 3, 4), \tag{22}$$

and the noise with the following distribution

$$f = \frac{1}{2}N(0, 20) + 20\frac{1}{\sqrt{2\sigma^2}}e^{-\sqrt{2/\sigma^2}|\epsilon|}, \tag{23}$$

where ϵ is a white noise uniformly distributed over the interval $[-0.5, 0.5]$ and $\sigma^2 = 0.25$. Fig.2(a) and Fig.2(b) display the noisy blur images. The LCLS fusion algorithm gives an average image estimate listed in Fig.2(c) and the wavelet fusion algorithm achieves the fused image shown in Fig.2(d). It can be seen that both the fused images have lost some contrast information. Figs.2 (e) and (f) give much better fused images by the L_1 estimation-based learning algorithm and the proposed learning algorithm with parameters $L = 255$, $N = 480$, $\alpha = 4.5$, and $\gamma = 1$, respectively. Table II lists computed results of the four algorithms. From Table II we see that the proposed learning algorithm gets the best results on image quality Q and QW. Moreover, the proposed learning algorithm is much faster than other three fusion algorithms. Finally, we use the CVX algorithm for the optimal fusion weight defined in (14). The CVX algorithm can give better result similar to the proposed learning algorithm but needs computation time being 3.797 second. Thus, the proposed learning algorithm also has a faster speed than the CVX algorithm.

In the following example, we study real images as test images.

TABLE II
PERFORMANCE COMPARISON OF FOUR ALGORITHMS IN EXAMPLE 2

Method	wavelet	LCLS	LAD	Proposed algorithm
CUP time	0.9782	0.8750	39.3750	0.4850
Q	0.4510	0.6348	0.9928	0.9930
QW	0.5648	0.6480	0.6530	0.6531

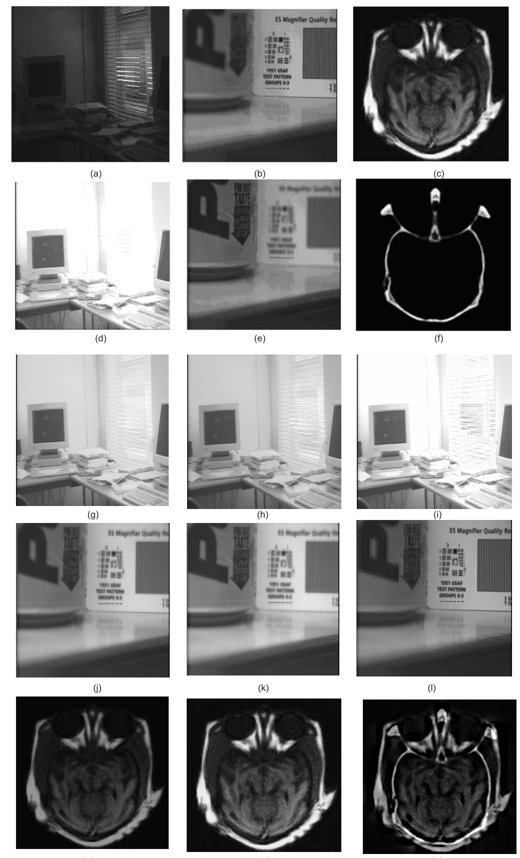

Fig. 3. Image fusion in example 3: (a)-(f) three pairs of input images; (g), (j), (m) fused images by the proposed fusion algorithm; (h), (k), (n) fused images by the L_1 estimation-based cooperative learning algorithm; (i), (l), (o) fused images by the wavelet algorithm.

TABLE III
PERFORMANCE COMPARISON OF THREE ALGORITHMS IN EXAMPLE 3

Case (a) and (d)	wavelet	LAD	Proposed algorithm
CPU time	0.65	35.56	0.171
Q_{NICE}	0.8051	0.8315	0.8298
Case (b) and (e)	wavelet	LAD	Proposed algorithm
CPU time	0.66	12.86	0.172
Q_{NICE}	0.8089	0.8049	0.830
Case (c) and (f)	wavelet	LAD	Proposed algorithm
CPU time	0.297	3.23	0.188
Q_{NICE}	0.8048	0.8472	0.830

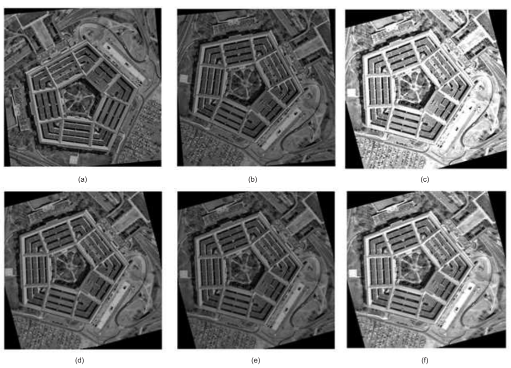

Fig. 4. Blind image fusion in example 4: (a)-(c) three aerial images; (d) and (e) registration images; (f) fused result of three registered images by the proposed fusion algorithm.

Example 3. Consider three pairs of real multisensor images obtained from the database of Manchester University. They are shown in Fig.3(a)-(f), respectively. To evaluate the performance of the proposed learning algorithm, we first compare it with the L_1 estimation-based cooperative learning (LAD) algorithm and the wavelet algorithm. Since the idea reference image is unknown in this example, we here use a nonlinear correlation information entropy (Q_{NICE}) approach which is viewed to be the most stable metric for image fusion algorithms [25]. Fig.3(g)-(o) display fused images by the three algorithms, respectively. As seen, the proposed learning algorithm can combine the complementary information from multiple sensors and improve the perceptual quality. Moreover, the proposed learning algorithm is much faster than the other algorithms, see the results listed in Table III. Next, we perform the CVX algorithm for the optimal fusion weight. Its computation time is 1.64 second, 1.671 second, and 1.375 second, respectively for the three pairs of real images. By contrast, we see that the proposed learning algorithm has also a faster speed than the CVX algorithm.

Example 4. Consider three aerial real images of city areas of China, shown in Fig.4(a)-(c). Let the image of Fig.4 (c) be a reference image. The three images are first registered by using the SIFT operator- based image registration algorithm [30] and we obtain two registered images with the reference image, shown in Fig.4(d) and (e), respectively. We then perform the proposed fusion algorithm to fuse three registered images listed in Fig.4 (c)-(e). Fig.4(f) displays the fused image, which improves percep- tual quality on the three registered images.

Finally, we give an example of signal fusion to illustrate some application limitation of the proposed fusion method.

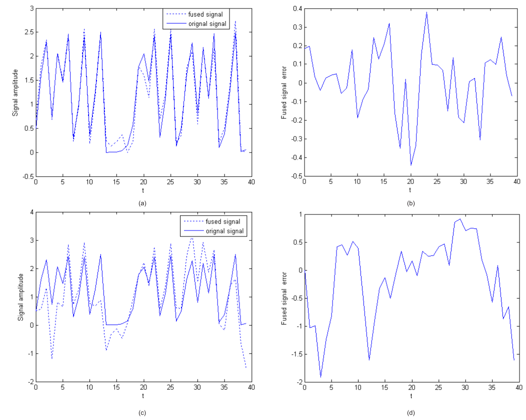

Fig. 5. Results of fused signal by the proposed fusion algorithm in example 5: (a) and (b) fused signal and its error in white Gaussian noise; (c) and (d) fused signal and its error in colored noise.

Example 5. Consider the signal sensor system given below:

$$\left\{ x_i(t) = 2.5 * s(t) + n_i(t), \quad i = 1, ..., K \right.$$

where $n_i(t)$ is measured noise, $s(t)$ is a chaotic signal processes generated by logistic mapping equation

$$v(t) = 4v(t-1)(1 - v(t-1)),$$

with the initial value $v(0) = 0.2$, 2.5 is a unknown scaling coefficient, K is the number of sensors. Let the number of measurements be $N = 40$ and the sensor number $K = 4$. Since $0 \le s(t) \le 1$, we take $\gamma_1 = 0$ and $\gamma_2 = 1$. For our test, we first consider one case: $n_i(t)$ is white Gaussian noise with variance being 0.5. Fig. 5(a) and Fig. 5(b) display the results of the fused signal based on the proposed fusion method. We next consider another case: the measurement noise is colored, which is modeled as a AR(1) process:

$$u(t) = 0.7u(t-1) + w(t)$$

where $w(t)$ is white Gaussian noise with variance being 0.5. Fig. 5(c) and Fig. 5(d) display the results of the fused signal based on the proposed fusion method. From computed results we see that the proposed fusion method can obtain better fusion results in white noise than the one in colored noise. This also implies an application limitation of the proposed fusion method in colored noise.

VI. CONCLUSION

An optimal fusion estimation method is developed for blind fusion of multiple sensor data and a fast learning algorithm is proposed. Compared with related fusion methods, the proposed approach is efficient when dealing with blind fusion under unknown scaling coefficients and arbitrary white noise field. Moreover, the proposed estimation method can be effectively used in the blind image fusion. Compared to the L_1 estimation-based cooperative learning algorithm, the proposed learning algorithm has a fast speed. Simulations also confirm the effectiveness of the proposed learning algorithm in the white noise case. Moreover, from computed results we see that the proposed fusion method can obtain better fusion results in white noise than the one in colored noise. This also implies an application limitation of the proposed fusion method in colored noise. Our next step is to improve the proposed fusion work to optimal fusion in the colored noise case and to extend the proposed fusion work to optimal fusion in the feature and decision levels. The applicability of the proposed fusion method in other areas is also our future study.

APPENDIX

THE PROOF OF THEOREM 1

Proof: Let the mapping $F(\mathbf{w}) = -(X^T X + \eta I)\mathbf{w} + X^T P_\Omega(X\mathbf{w})$ and let $\mathbf{w}^*$ be a solution satisfying (14). By the KKT condition we know that $\mathbf{w}^*$ satisfies $-(X^T X + \eta I)\mathbf{w}^* + X^T P_\Omega(X\mathbf{w}^*) = 0$ and thus it is an equilibrium point of (18a). On the other hand, we define a Lyapunov function

$$V(\mathbf{w}) = \frac{1}{2}\|\mathbf{w} - \mathbf{w}^*\|_2^2.$$

Note that $F(\mathbf{w}^*) = 0$. Then $dV(\mathbf{w})/dt = (\mathbf{w} - \mathbf{w}^*)^T \frac{d\mathbf{w}}{dt} = \lambda(\mathbf{w} - \mathbf{w}^*)^T(F(\mathbf{w}) - F(\mathbf{w}^*))$. One one side, $(F(\mathbf{w}) - F(\mathbf{w}^*))^T(\mathbf{w} - \mathbf{w}^*) = -(\mathbf{w} - \mathbf{w}^*)^T(X^T X + \eta I)(\mathbf{w} - \mathbf{w}^*) + [X^T P_\Omega(X\mathbf{w}) - X^T P_\Omega(X\mathbf{w}^*)]^T(\mathbf{w} - \mathbf{w}^*)$. On the other side, $[X^T P_\Omega(X\mathbf{w}) - X^T P_\Omega(X\mathbf{w}^*)]^T(\mathbf{w} - \mathbf{w}^*) = [P_\Omega(X\mathbf{w}) - P_\Omega(X\mathbf{w}^*)]^T X(\mathbf{w} - \mathbf{w}^*) \le \|X(\mathbf{w} - \mathbf{w}^*)\|_2^2$. Moreover, $(\mathbf{w} - \mathbf{w}^*)^T(X^T X + \eta I)(\mathbf{w} - \mathbf{w}^*) = \|X(\mathbf{w} - \mathbf{w}^*)\|_2^2 + \eta\|\mathbf{w} - \mathbf{w}^*\|_2^2$. Then $(F(\mathbf{w}) - F(\mathbf{w}^*))^T(\mathbf{w} - \mathbf{w}^*) \le -\eta\|\mathbf{w} - \mathbf{w}^*\|_2^2$. It follows that $dV(\mathbf{w})/dt = \lambda(F(\mathbf{w}) - F(\mathbf{w}^*))^T(\mathbf{w} - \mathbf{w}^*) \le -2\lambda\eta V(\mathbf{w})$. Then $V(\mathbf{w}(t)) = V(\mathbf{w}(t_0))e^{-2\lambda\eta(t-t_0)}$. Thus we have

$$\|\mathbf{w}(t) - \mathbf{w}^*\|_2^2 = 2V(\mathbf{w}(t_0))e^{-2\lambda\eta(t-t_0)}.$$

Furthermore, we have $\|\mathbf{z}(t) - \mathbf{z}^*\|_2^2 = \|g(X\mathbf{w}(t)) - g(X\mathbf{w}^*)\|_2^2 \le \|X\|_2^2\|\mathbf{w}(t) - \mathbf{w}^*\|_2^2 = 2\|X\|_2^2 V(\mathbf{w}(t_0))e^{-2\lambda\eta(t-t_0)}$. It follows the conclusion of Theorem 1.

REFERENCES

[1] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proc. IEEE*, vol. 85, no. 1, pp. 6–23, Jan. 1997.

[2] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. New York, NY, USA: Wiley, 2003.

[3] S. Duncan and S. Sameer, "Approaches to multisensor data fusion in target tracking: A survey," *IEEE Trans. Knowl. Data Eng.*, vol. 18, no. 12, pp. 1696–1710, Dec. 2006.

[4] R. C. Luo Ren, C. C. Chia, and L. C. Chi, "Multisensor fusion and integration: Theories, applications, and its perspectives," *IEEE Sensors J.*, vol. 11, no. 12, pp. 3122–3138, Dec. 2011.

[5] D. P. Mandic, M. Golz, A. Kuh, D. Obradovic, and T. Tanaka, *Signal Processing Techniques for Knowledge Extraction and Information Fusion*, New York, NY, USA: Springer-Verlag, 2008.

[6] H. F. Durrant-Whyte, "Integrating distributed sensor information: An application to robot system coordinate," *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, 1985, pp. 415–419.

[7] Y. Zhou and H. Leung, "A maximum likelihood approach for multisensor data fusion applications," *Proc. SPIE*, vol. 3376, pp. 186–195, Apr. 1998.

[8] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graph. Models Image Process.*, vol. 57, no. 3, pp. 235–245, 1995.

[9] R. S. Blum, "On multisensor image fusion performance limits from an estimation theory perspective," *Inf. Fusion*, vol. 7, no. 3, pp. 250–263, 2006.

[10] Y. S. Xia, H. Leung, and E. Bossé, "Neural data fusion algorithms based on a linearly constrained least square method," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 320–329, Mar. 2002.

[11] S. Suranthiran and S. Jayasuriya, "Optimal fusion of multiple nonlinear sensor data," *IEEE Sensors J.*, vol. 4, no. 5, pp. 651–660, Oct. 2004.

[12] Y. S. Xia and M. Kamel, "Novel cooperative neural fusion algorithms for image restoration and image fusion," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 367–381, Feb. 2007.

[13] M. Kumar and S. Dass, "A total variation-based algorithm for pixel-level image fusion," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 2137–2143, Sep. 2009.

[14] M. Xu, H. Chen, and P. K. Varshney, "An image fusion approach based on Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 5116–5127, Dec. 2011.

[15] Y. S. Xia and M. S. Kamel, "Cooperative learning algorithms for data fusion using novel L_1 estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1083–1095, Mar. 2008.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[17] M. Grant, S. Boyd, and Y. Y. Ye, "Disciplined convex programming," in *Global Optimization: From Theory to Implementation*, L. Liberti and N. Maculan, Eds. Norwell, MA, USA: Kluwer, 2005.

[18] D. Mandic and J. A. Chambers, "Recurrent neural networks for prediction: Learning algorithms, architectures and stability," in *Adaptive and Learning Systems for Signal Processing, Communications and Control*. New York, NY, USA: Wiley, 2001.

[19] Y. S. Xia, "An extended projection neural network for constrained optimization," *Neural Comput.*, vol. 16, no. 4, pp. 863–883, 2004.

[20] Y. S. Xia, "A compact cooperative recurrent neural network for computing general constrained L_1 norm estimators," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3693–3697, Sep. 2009.

[21] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, 2006.

[22] A. Cichocki and R. Unbehauen, *Neural Networks for Optimization and Signal Processing*. New York, NY, USA: Wiley, Nov. 1994.

[23] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[24] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. ICIP*, Barcelona, Spain, 2003, pp. 173–176.

[25] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiere, and W. Wu, "Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 94–109, Jan. 2012.

[26] T. H. Corman, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*. New York, NY, USA: McGraw-Hill, 1990, ch. 2.

[27] B. Khaleghi, A. Khamis, and F. O. Karray, "Multisensor data fusion: A review of the state-of-the-art," *Inf. Fusion*, vol. 14, no. 1, pp. 28–44, 2013.

[28] Z. L. Deng, Z. Peng, W. Qi, J. Liu, and Y. Gao, "Sequential covariance intersection fusion Kalman filter," *Inf. Sci.*, vol. 189, pp. 293–309, Apr. 2012.

[29] J. Hu, L. Xie, and C. Zhang, "Diffusion Kalman filtering based on covariance intersection," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 891–902, Feb. 2012.

[30] H. Wen and Y. S. Xia, "An improved SIFT operator-based image registration using cross-correlation coefficient," in *Proc. Int. Congr. Image Signal Process.*, Oct. 2011, pp. 869–873.

Youshen Xia received the B.S. and M.S. degrees in computational mathematics from Nanjing University, Nanjing, China, in 1982 and 1989, respectively, and the Ph.D. degree in computational intelligence from The Chinese University of Hong Kong, Hong Kong, in 2000. He is currently a Professor with Fuzhou University, China and his research interests include design and analysis of neural dynamical optimization approaches with applications to signal and image processing, data fusion, and system identification.

Henry Leung (M'90) received the Ph.D. degree in electrical and computer engineering from McMaster University, Hamilton, ON, Canada.

He is currently a Professor with the Department of Electrical and Computer Engineering, University of Calgary, Calgary, AB, Canada. He was with Defense Research Establishment, Ottawa, ON, Canada, where he was involved in the design of automated systems for air and maritime multisensor surveillance. His current research interests include chaos, computational intelligence, information fusion, multimedia, robotics, sensor networks, and wireless communications.