EFFICIENT BLIND SEPARATION OF REFLECTION LAYERS WITH NONPARAMETRIC TRANSFORMATIONS

Han Li⋆ *Kun Gai*†⋆ *Pinghua Gong*⋆ *Changshui Zhang*⋆

⋆ State Key Laboratory on Intelligent Technology and Systems,
Tsinghua National Laboratory for Information Science and Technology (TNList),
Department of Automation, Tsinghua University, Beijing, 100084, P.R.China
† Alibaba Group, Beijing, 100022, P.R.China

ABSTRACT

Superimposed images are very common when taking photos behind glass. We address the reflection separation problem using multiple superimposed images photographed in different viewpoints. With viewpoints changing, the reflected scenes could contain arbitrarily complicated variations between mixtures, like human's motions or other nonrigid motions. In this article, we propose a moderate hypothesis to tackle the reflected scenes' arbitrary variations as well as the parametric transformations of transmitted scenes. To rapidly recover high-quality image layers, we propose an Efficient Superimposition Recovering Algorithm (ESRA) by extending the framework of accelerated gradient method. Our recovering method has good converging performance and is more than 30 times faster than state-of-the-art methods. Experimental results on synthetic and real world images demonstrate that our method is promising.

***Index Terms*—** Blind separation, nonparametric transformation, optimization

1. INTRODUCTION

Superimposed images with undesired reflection scenes are often obtained when people take photos behind glasses. These superimposed images are linear mixtures of two transparent image layers [1, 2, 3, 4], one of which is the transmitted scene of interest, the other is the reflected scene. These superimpositions seriously disturb the viewing of the layers of interest. Thus separating superimposed images and recovering all layers are needed.

To achieve automatic and good separation, two or more different superimposed images are needed. A convenient way to obtain multiple photos is to change viewpoints and take different photos (see Fig. 1). While by changing viewpoints, two transparent layers are no longer static mixing, ICA like methods [1, 5, 6, 7, 8, 9, 10, 11] are not applicable. Although several technologies have recently been developed to separate superimposed images with different variations: 1) with parametric motions [2, 3, 12, 13], 2) with time/position variant mixing [14, 15]. They assume the variations of all layers conform to the same parametric model. In Fig. 1, the transmitted layer is a planar painting, and its variations conform to certain rules and can be modeled with parametric transformations, like planar transformations. However, transformations of reflected layers contain complex variations, like human motions (see Fig. 1), which are hard to be modeled in a parametric way. Gai et al. [16] give an example of handling this more common problem with their parametric method. They use different parametric layers to model the reflected layers with different appearances (in different snapshots) and assume the reflected layers are totally independent of each other. However, the reflected layers could be arbitrarily similar, making their assumption unapplicable in some cases. More reasonable hypothesis for this general blind separation problem and theoretic study are needed.

Fig. 1. Real superimposed photos of a painting in glass frame.

Moreover, when the variations of latent transparent layers are well estimated, fast recovering high quality latent layers is preferred. Existing methods [4, 17] can fast restore the transparent layers, but they can not ensure good recovery and have color-bias problem (as illustrated in [16]). Methods used in [2, 13, 16] can recover latent layers of high quality, but they are slow (see Table 1).

In this article, we only assume the parametric transformations for transmitted layers, and formulate this more common superimposition problem in detail. We propose a more moderate hypothesis to handle the complicated variations of the reflected layers and give a theoretic study about our blind estimation method. Moreover, we develop an Efficient Superimposition Recovering Algorithm (ESRA) by extending the framework of accelerated gradient method [18, 19]. In the proposed ESRA, a key building block (in each iteration) is the proximal operator calculating, here we propose a Parallel Algorithm with Constrained Total Variation (PACTV) method to efficiently find the optimal solution in every iteration. Our recovering method not only reconstructs high-quality layers without color-bias problem, but also theoretically guarantees good convergence performance.

2. PROBLEM FORMULATION AND HYPOTHESES

As discussed in Section 1, we release the parametric transformation assumption about reflected layers and give our more general formulation. We consider the captured m superimposed images I_i ($i = 1, \cdots, m$) are linear mixtures of the transmitted layer L^t and the reflected layer L^r as in many other work (e.g. [1, 2, 3, 4]). We assume the variations of transmitted layers conform to a user specified parametric transformations $f(x, \theta_i)$ with unknown parameter θ_i, here $x = (x_1, x_2)^\top$ is a 2D vector that represents the pixel coordinates. In practical situations, reflected layers between mixtures contain nonrigid motions, like human's motions , or occlusions. Therefore, we formulate the complicated variations of reflected layers with a nonparametric transformation $u_i(x)$, i.e $L^{r(i)}(x) = L^r(u_i(x))$. Here $L^{r(i)}$ is the reflected layer in ith mixture, and L^r is the referenced reflected image scene. Besides, due to different lighting conditions or reflection angles, layers' overall intensity may change from one mixture to another, leading to different mixing coefficients a_{ij}. Therefore, we finally formulate m superimposed images as

$$I_i(x) = a_{i1}L^t(f(x,\theta_i)) + a_{i2}L^{r(i)}(x), \quad \text{s.t } a_{i1}, a_{i2} \geq 0, \; i = 1, \cdots, m, \tag{1}$$

where $I_i \in \mathbb{R}^{h \times w}$ is the ith superimposed image, which is called the ith *mixture*. $x = [x_1, x_2]^T$ represents the pixel coordinates, and a_{ij} is the nonnegative mixing coefficient. Without loss of generality, we regard layers in the first mixture as the reference layers, so $a_{11} = a_{12} = 1$ and $f(x, \theta_1) = u_1(x) = x$. In Fig 2, the four mixtures (Fig. 2(a)) are superimposed by Lena image and another image using mixing model (1) with different transformation parameters θ_i and mixing coefficients a_{i1}, a_{i2}, respectively.

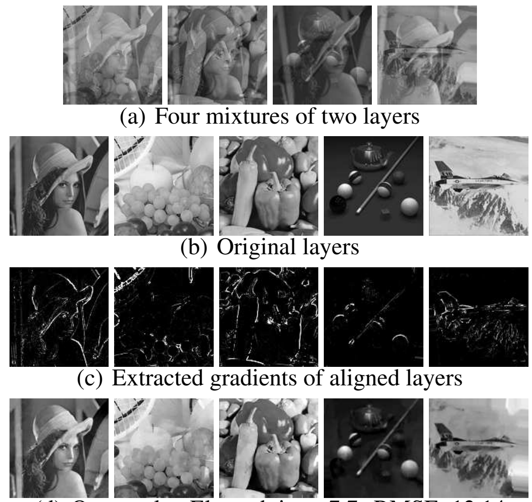
With mixing model (Eq. 1), only the *mixtures* I_i are known as input. Both the latent layers $L^t, L^{r(i)}$ and mixing coefficients a_{i1} and transformation parameters θ_i are unknown. It makes our problem very challenging. To solve this complicated problem, we review the previous hypotheses in [16] and propose a new hypothesis for more general cases. Given parametric transformations $f(x, \theta_i)$, we first introduce denotations about image layer gradients: $\nabla L\big(f(x,\theta_i)\big) = \frac{dL(y)}{dy}\Big|_{y=f(x,\theta_i)}$ and $D\big(L(f(x,\theta_i))\big) = \frac{df^\top(x,\theta_i)}{dx}\nabla L\big(f(x,\theta_i)\big)$. Here $\nabla = ((\partial)/(\partial x_1), (\partial)/(\partial x_2))^\top$ is the gradient operator, and L represents the transparent layer L^t or $L^{r(i)}$. We follow hypotheses in [16] about natural image gradients' statistical properties: 1) **Sparsity**: the gradients of transparent layer ∇L are sparse; 2) **Non-correlation**: if the parametric layers in two different mixtures are not aligned, the correlations between their gradients are nearly zero; 3) **Independence between different transparent layers**: the gradients from any two different transparent layers are totally independent, leading to very small correlations between their gradients.

These three hypotheses aim to solve the parametric superimposition model in [16], which is not enough here. For more general cases, the variations of reflected layers contain nonrigid motions and cannot be modeled with a uniform parametric transformation. To handle such reflected layers with different appearances, Gai et al. [16] treat reflected layers in different images as different layers, and assume they are independent of each other. However, the reflected scenes could be arbitrarily similar. For example, as shown in Fig. 4, the reflected layers of synthetic mixtures all have the same part and are not independent. Thus the independence hypothesis among them is not reasonable. Therefore, we propose a moderate but reasonable hypothesis to tackle more general cases about reflected layers' variations, including the nonindependent cases.

Hypothesis 1. ***Nonparametric transformations of the reflected layers:*** *for any i, given transformation parameter θ_i of transmitted layer, we have:*

$$\mathbb{E}\Big[\big\langle a_{i1}DL^t(f(x,\theta_i)), DL^t(f(x,\theta_i))\big\rangle\Big] > \max_s \mathbb{E}\Big[\big\langle a_{i2}DL^{r(1)}(f(x,s)), \nabla L^{r(i)}(x)\big\rangle\Big]. \tag{2}$$

Here $\mathbb{E}[\cdot]$ is the expectation w.r.t. x, and $\langle \cdot, \cdot \rangle$ represents the inner product of two vectors. Gai et al.'s independence hypothesis [16] means that $\mathbb{E}\Big[\Big\langle a_{i2}DL^{r(1)}(f(x,s)), \nabla L^{r(i)}(x)\Big\rangle\Big] \approx 0$, given any parameter s. However, the mixtures shown in Fig. 4 violate this implication. Here we only assume the maximum correlation between any two reflected layers gradients is upper bounded. Note that the left term of (2) is a positive definite form and will be significantly larger than zero. It is a moderate condition for more general cases (including the mixtures in Fig. 4). On the other hand, if our hypothesis is violated, it implies that the variations of the reflected layers can be modeled with parametric transformations to some extent (the correlation is large under a particular transformation parameter). Methods in [13, 16] have already solved such rare cases.



(a) Four mixtures of two layers

(b) Original layers

(c) Extracted gradients of aligned layers

(d) Our results. Elapsed time: 7.7s RMSE: 12.14

Fig. 2. Demonstration of mixtures and their gradients.

3. BLIND ESTIMATION

In this section, we estimate the transmitted layers' transformation parameters in the gradient domain with the hypotheses mentioned above. Then the unknown mixing coefficients a_{i1} are estimated by the cluster method [2, 13, 16].

To search for the transmitted layer's transformation in the second mixture, we warp the first mixture I_1 by a parametric transformation $f(\cdot, s)$ with a searching parameter s, and match its gradients with the second mixture's gradients ∇I_2. Consider the following transformation objective function of these two mixtures:

$$O(s) = \mathbb{E}\left[\left\langle D\left(I_1\big(f(x,s)\big)\right), \nabla I_2(x)\right\rangle\right]. \tag{3}$$

By use of the mixing model (1) and definitions of $D\left(I_1\big(f(x,s)\big)\right)$, $\nabla I_2(x)$, the objective function (3) can be expanded as

$$O(s) = C^t(s) + C^r(s). \tag{4}$$

Here $C^t(s), C^r(s)$ are the correlation functions of the transmitted layers and reflected layers, respectively. We ignore the correlation between transmitted and reflected layers due to the independence hypothesis. $C^t(s)$ and $C^r(s)$ can be written as follows:

$$C^t(s) = a_{21}\mathbb{E}\Big[\nabla^\top L^t\big(f(x,s)\big)\frac{df(x,s)}{dx^\top}\frac{df^\top(x,\theta_2)}{dx}\nabla L^t\big(f(x,\theta_2)\big)\Big], \tag{5}$$

$$C^r(s) = a_{22}\mathbb{E}\Big[\nabla^\top L^{r(1)}\big(f(x,s)\big)\frac{df(x,s)}{dx^\top}\nabla L^{r(2)}\big(x\big)\Big]. \tag{6}$$

Here we assume $f(x,s)$ is an almost injective mapping, i.e, for $C^t(s)$, if $s \neq \theta_2$, we have $f(x,s) \neq f(x,\theta_2)$ for most x. Then due to the noncorrelation hypothesis, $\nabla L^t(f(x,s))$ and $\nabla L^t(f(x,\theta_2))$ is uncorrelated, thus $C^t(s) \approx 0$. Otherwise, if $s = \theta_2$, $C^t(s)$ turns to a positive definite form and will be significantly larger than zero. In all, we get $C^t(s) = C^t(\theta_2) > 0$, if $s = \theta_2$, or $C^t(s) = 0$, otherwise.

As illustrated in [13], if all layers conform to a parametric model, θ_2 can be found by maximizing (3). **While for our more general cases, can we still find θ_2 by maximizing (3)**? Gai et al. [16] assume all the reflected layers belong to different independent parametric layers, so $C^r(s) = 0, \; \forall\, s$. However, we find that if $\max_s C^r(s) < C^t(\theta_2)$, we can still find θ_2 in the same way. Because given any $s \neq \theta_2$, the following inequalities hold:

$$O(s) = C^r(s) < C^t(\theta_2) \leq C^t(\theta_2) + C^r(\theta_2) = O(\theta_2). \quad (7)$$

The first inequality is indeed our nonparametric transformations hypothesis (2), which is a more moderate condition to guarantee the blind estimation method work. Therefore, by our theoretic analysis, even though $C^r(s)$ may be larger than zero with some s, maximizing $O(s)$ can still find the optimal solution θ_2.

In this work, we use planar transformations [13] to describe the variations of transmitted layers. Note, any other form of almost injective parametric transformation can also be applied here. By maximizing $O(s)$ with discrete grid search method [13, 16], the transformation parameter θ_2 is obtained. And by exchanging other *mixtures* I_i ($i = 3, ..., m$) for I_2, other transformation parameters θ_i are obtained. If the transformation parameters θ_i are well-estimated, we can employ the general cluster method [2, 13, 16] to estimate the unknown mixing coefficients of transmitted layers and assign the significant gradients in $\nabla_k I_1(x)$ to latent different transparent layers. We use $E^t, E^{r(i)}$ denote the extracted gradients of $L^t, L^{r(i)}$, respectively. Fig. 2(c) shows the extracted gradients of latent transmitted layers in Fig. 2(a).

4. FAST RECONSTRUCTION OF LAYERS

With the extracted gradients of each layer, the reconstruction step is the final crucial part for reflection separation. In this section, we propose an Efficient Superimposition Recovering Algorithm (ESRA) to fast recover the high quality latent layers.

4.1. Efficient Superimposition Recovering Algorithm

With estimated parameters θ_i, we align the transmitted layers by warping *mixtures* I_i with $f^{-1}(x, \theta_i)$. Then our mixing model is rewritten as:

$$I_i(f_i^{-1}(x)) = a_{i1}L^t(x) + a_{i2}L^{r(i)}(f_i^{-1}(x)), \qquad i = 1, \cdots, m. \quad (8)$$

For simplicity, we use $I_i(x)$ to represent $I_i(f_i^{-1}(x))$. $L_1(x)$ and $L_{i+1}(x)$ denote $L^t(x)$ and $a_{2i}L^{r(i)}(f_i^{-1}(x))$, respectively. Let $E^i(x)$ stand for the extracted gradients from $L^i(x)$. To recover high quality latent image layers, we propose to employ L_1 penalty on the extracted gradients and nonnegative constraints on the layers' intensities along with the L_2 loss of the mixing model. Thus our recovering objective function is written as:

$$\begin{aligned} \min_{0 \leq l^{vec} \leq 1} F(l^{vec}) = \lambda \sum_{x,i=1}^{m+1} |\nabla L^i(x) - E^i(x)| \\ + \sum_{x,i=1}^{m} \frac{1}{2} \big(I_i(x) - a_{i1}L_1(x) - L_{i+1}(x)\big)^2 \end{aligned} \quad (9)$$

where $l^{vec} \doteq [\mathrm{vec}^\top(L^1), \cdots, \mathrm{vec}^\top(L^{m+1})]^\top$ is a large vector containing all pixel values in all latent layers. The first L_1 term enforces the agreement between reconstructed layer gradients and extracted layer gradients, while the second L_2 term tends to satisfy our mixing mode. Since the extracted gradients are nonzero at very few coordinates, the L_1 penalty not only prefers layers with sparse gradients but also avoids over-smooth results. λ is a trade off coefficient.

To solve the nonsmooth convex optimization model (9) efficiently, we denote

$$\begin{aligned} f(l^{vec}) &= \sum_{x,i=1}^{m} \frac{1}{2}\big(I_i(x) - a_{i1}L_1(x) - L_{i+1}(x)\big)^2, \ \text{s.t } 0 \leq l^{vec} \leq 1, \\ g(l^{vec}) &= \lambda \sum_{x,i=1}^{m+1} |\nabla L^i(x) - E^i(x)|. \end{aligned} \quad (10)$$

Here $g(l^{vec})$ is the ℓ_1 penalty on the extracted gradients and $f(l^{vec})$ corresponds to the L_2 loss and nonnegative constraints. Note that $f(l^{vec})$ is continuously differentiable, of which Lipschitz constant $L(f) = \sum_i a_{i1}^2 + 1$. We note the objective function in (9) is a composite function of a differential term $f(l^{vec})$ and a non-differential term $g(l^{vec})$. Denote

$$P_{L_s, l_{k-1}^{vec}}(l^{vec}) = f(l_{k-1}^{vec}) + \langle \nabla f(l_{k-1}^{vec}), l^{vec} - l_{k-1}^{vec} \rangle + \frac{L_s}{2}\|l^{vec} - l_{k-1}^{vec}\|^2,$$

which is the first order Taylor expansion of $f(l^{vec})$ at l_{k-1}^{vec}, with the squared Euclidean distance between l^{vec} and l_{k-1}^{vec} as the regularization term. The traditional gradient descent algorithm obtains the solution at the k-th iteration ($k \geq 1$) by $l_k^{vec} = \arg\min P_{L_s, l_{k-1}^{vec}}(l^{vec}) + g(l^{vec})$ with a proper step size L_s (greater than $L(f)$). Here we propose to employ the accelerated gradient descent [18, 19] to solve the reconstruction problem, named Efficient Superimposition Recovering Algorithm (ESRA). Here we generate a solution at the k-th iteration ($k \geq 1$) by computing the following proximal operator

$$l_k^{vec} \to \arg\min_{0 \leq l^{vec} \leq 1} P_{L_s, Y_k}(l^{vec}) + g(l^{vec}) \quad (11)$$

where $Y_1 = l_0^{vec}$ and $Y_k = l_{k-1}^{vec} + \frac{t_{k-2}-1}{t_{k-1}}(l_{k-1}^{vec} - l_{k-2}^{vec})$ for $k \geq 1$. We note that Y_k is a linear combination of l_{k-1}^{vec} and l_{k-2}^{vec}. The combination coefficient plays an important role in the convergence of the algorithm. As suggested by [20], we set $t_0 = 1$ and $t_k = (1 + \sqrt{t_{k-1}^2 + 1})/2$ for $k \geq 1$. According to the theoretical analysis in [20], this accelerated gradient descent method can get within $O(1/k^2)$ of the optimal objective value after k steps. While solving problem (11) is still very challenging, we propose a Parallel Algorithm with Constrained Total Variation (PACTV) method to find the optimal solution, which is presented in the sequel.

4.2. PACTV via dual approach

Given problem (11), we observe it can be solved block separable in the following way. If we denote $Y_k - \frac{1}{L_s}\nabla f(Y_k) \doteq [\mathrm{vec}^\top(d_1), \cdots, \mathrm{vec}^\top(d_{m+1})]^\top$ ($d_i \in \mathbb{R}^{h \times w}$ $i = 1, \cdots, m+1$), we can split $Y_k - \frac{1}{L_s}\nabla f(Y_k)$ into $m+1$ separable parts. Then by employing the definition of (10), we transform (11) as follows:

$$l_k^{vec} = \operatorname*{argmin}_{0 \leq l_k^{vec} \leq 1} \Big\{ \sum_{x,i=1}^{m+1} \big(\lambda|\nabla L^i(x) - E^i(x)| + \frac{L_s}{2}||L^i(x) - d_i(x)||^2\big) \Big\}.$$

As illustrated above, finding l_k^{vec} is to solve following $m+1$ separable problems with constrained total variation in parallel:

$$\min_{0 \leq L \leq 1} \sum_x \big(\frac{1}{2}||L(x) - d(x)||^2 + \beta|\nabla L(x) - E(x)|\big). \quad (12)$$

Here $\beta = \lambda/L_s$, and L, d, E represent L^i, d_i, E^i, respectively. Similar with the image denoising problem [20, 21] (note we have extra constants $E(x)$ in our objective function (12)), we propose a dual approach to solve (12). More details can be found in [22]. Our key idea is using Fast Gradient Projection Method (FGP) [21] to solve the $m+1$ dual problems of (12) in parallel, which is called Parallel Algorithm with Constrained Total Variation (PACTV). The computation complexity in each iteration of FGP is $O(hw)$ and the converging rate of FGP is also $O(1/k^2)$. Finally, we couple the optimal L_i^* ($i = 1, \ldots, m+1$) together, and resize them into a vectorial form to achieve l_k^{vec}.

In our implementations, we set $\lambda = 0.01$ (Fig. 5(b) with a larger $\lambda = 0.12$), and we also set $L_s = 2L(f)$ to ensure a constant stepsize. And initial value of l^{vec} is zero. The final recovered reflected layers of (9) should be warped with f_i and be enhanced to be visible. Our experiments illustrate that 100 iterations in ESRA is enough to achieve the satisfied results. What's more, our recovering method launches a general optimization framework and can be extended to solve other reconstruction problems in [13, 16].

(a) Input | (b) ESRA 124.2s | (c) SPBSM[16] 6690.6s | (d) IRLS[4] 238.8s | (e) Weiss[17]

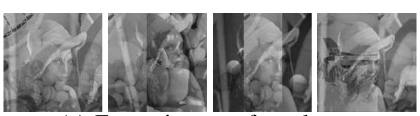
Fig. 3. Separation results for real photos shown. The elapsed time of the first three method are listed.

5. EXPERIMENTS

In this section, we show experiments of our algorithm on both synthetic and real superimposed images. Since other separation methods can not be directly applied for our superimposed images, we only compare the final recovered methods based on our estimated parametric transformations and mixing coefficients. To illustrate advantages of our approach, our recovered results are compared with other reconstruction methods, like IRLS [4], SPBSM [16], as well as Weiss [17] method. The settings of SPBSM and IRLS are the same with in [4, 16], respectively. For color images, we use their grayscales to estimate and then reconstruct R,G,B channels separately. We run all experiments on an Intel i5 PC (3.3GHz CPU, 8GB RAM) with MATLAB implementation.

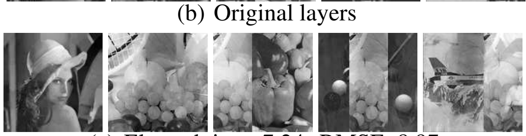
The simulation results are shown in Fig. 2 and Fig. 4. Each mixture is superimposed by the Lena layer and another layer with different mixing coefficients and planar transformations. The recovered results are given in Fig. 2(d) and Fig. 4(c), which are high quality with a relative short computation time. Figure 4 gives a special case when partial reflected scenes are the same between all mixtures. The transmitted layers' correlation C^t is 18.4, while the maximum correlation C^r between any two reflected layers is 7.11 and cannot be ignored by the independence hypothesis in [16]. However, our nonparametric hypothesis (2) works for this special case, and our approach also give the good separation results (shown in Fig. 4(c)).

Figure 3 shows the separation results of different methods on real superimposed images in Fig. 1. Since blind separation method in [4] is not automatically and needs hundreds of labeled layers'



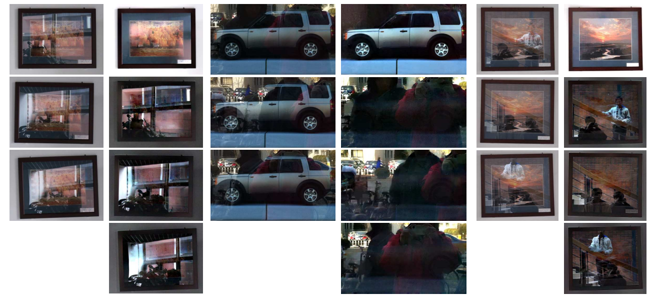
(a) Four mixtures of two layers

(b) Original layers



(c) Elapsed time: 7.24s RMSE: 8.97

Fig. 4. Separation results for a synthetic example when partial scenes are the same in all reflected layers.



(a) (b) (c)

Fig. 5. More real mixtures ((c) are supplied by Gai et al. [16]).

edges, we first align the transmitted layers of mixtures and then use the locations of our extracted layers' gradients as input labeled edges. After 15 IRLS iterations, the recovered transmitted layer is normalized and other reflected layers are achieved by Eq. (1). After aligned the transmitted layers, Sarel and Irani [6] propose to recover latent layers with Weiss' approach [17]. Compared with SPBSM method, our method also achieves good separation results, which is nearly 35 times faster. Since IRLS [4] can not guarantee convergence and does not consider constraints on pixel values, the recovered results are poor. Sarel and Irani [6] only use their estimated gradients to recover layers' intensities, their results are sensitive with gradients errors and have color-bias problem.

Moreover, our approach also gives good performance on other real superimposed images, shown in Fig. 5. In addition, when the transmitted layers' variations do not perfectly conform to planar transformations in real practice (see Fig. 3,5), our method still gives good results. It illustrates that our method is robust even if the real 3D variations are just approximately modeled by planar transformations. Note that besides planar transformation, any other form of parametric transformations can also be applied in our blind separation method. The computation time of different methods are given in Table 1. As is shown, although IRLS runs faster when dealing with small size mixtures, they obtain poor results. SPBSM can achieve high quality results. However, its computation time scales poorly as the number of unknowns increases. Above all, our recovering method can give good results with a short computation time.

6. CONCLUSION

In this paper, we address the reflection separation problem from more general superimposed images with different viewpoints, of which the reflected layers have complicated variations and cannot be modeled with a uniform parametric transformation. We develop an efficiently reconstructing method named ESRA and propose a new hypothesis to theoretically support the blind estimation step. The experiments on real and synthetic mixtures are promising.

7. ACKNOWLEDGEMENTS

Supported by 973 Program (2013CB329503), NSFC (Grant No. 91120301), Beijing Municipal Education Commission Science and Technology Development Plan Key Project (No. KZ201210005007).

Mixtures	Unknowns	ESRA	IRLS[4]	SPBSM[16]
Figure 2(a)	82k pixels	7.7s	**7.0s**	378.8s
Figure 4(a)	82k pixels	**7.2**s	**7.2s**	575.7s
Figure 5(b)	307k pixels	40.6s	**39.7s**	1435.2s
Figure 1	420k pixels	**40.9s**	61.9s	1530.3s
Figure 5(a)	480k pixels	**43.7s**	68.6s	2137.3s
Figure 5(c)	480k pixels	**41.4s**	79.6s	2230.2ss

Table 1. The average single-channel recovering time of different recovering methods for different-sized mixtures(1k = 1000). Fig. 5(b) uses a larger λ. It has a smaller stepsize in FGP and takes more time.

8. REFERENCES

[1] Alexander M. Bronstein, Michael M. Bronstein, Michael Zibulevsky, and Yehoshua Y. Zeevi, "Sparse ica for blind separation of transmitted and reflected images," *Int'l J. Imaging Science and Technology*, vol. 15, no. 1, pp. 84–91, 2005.

[2] Kun Gai, Zhenwei Shi, and Changshui Zhang, "Blindly separating mixtures of multiple layers with spatial shifts," in *Proc. CVPR*, 2008, pp. 1–8.

[3] Richard Szeliski, Shai Avidan, and P. Anandan, "Layer extraction from multiple images containing reflections and transparency," in *Proc. CVPR*, 2000, pp. 246–253.

[4] Anat Levin and Yair Weiss, "User assisted separation of reflections from a single image using a sparsity prior," *TPAMI*, vol. 29, no. 9, pp. 1647–1654, 2007.

[5] Hany Farid and Edward H. Adelson, "Separating reflections and lighting using independent components analysis," in *Proc. CVPR*, 1999, vol. 1, pp. 262–267.

[6] B. Sarel and M. Irani, "Separating transparent layers through layer information exchange," in *Proc. ECCV*, 2004, pp. 328–341.

[7] Yoav Y. Schechner, Joseph Shamir, and Nahum Kiryati, "Polarization-based decorrelation of transparent layers: The inclination angle of an invisible surface," in *Proc. ICCV*, 1999, pp. 814–819.

[8] Y. Y. Schechner, J. Shamir, and N. Kiryati, "Polarization and statistical analysis of scenes containing a semireflector," *J. Opt. Soc. Amer. A*, vol. 17, no. 1-3, pp. 276–284, May 2000.

[9] N. Kong, YW Tai, and SY Shin, "High-quality reflection separation using polarized images," *TIP*, , no. 99, pp. 1–1, 2011.

[10] Y. Y. Schechner, N. Kiryati, and R. Basri, "Separation of transparent layers using focus," in *Proc. ICCV*, 1998, pp. 1061–1066.

[11] Yoav Y. Schechner, Joseph Shamir, and Nahum Kiryati, "Blind recovery of transparent and semireflected scenes," in *Proc. CVPR*, 2000, pp. 38–43.

[12] Efrat Be'ery and Arie Yeredor, "Blind separation of superimposed shifted images using parameterized joint diagonalization," *TIP*, vol. 17, no. 3, pp. 340–353, March 2008.

[13] Kun Gai, Zhenwei Shi, and Changshui Zhang, "Blind separation of superimposed images with unknown motions," in *Proc. CVPR*, 2009, pp. 1881–8.

[14] R. Kaftory and Y.Y. Zeevi, "Blind separation of images obtained by spatially-varying mixing system," in *ICIP*. IEEE, 2008, pp. 2604–2607.

[15] R. Kaftory and Y.Y. Zeevi, "Blind separation of position varying mixed images," in *ICIP*, 2009, pp. 3913–3916.

[16] K. Gai, Z. Shi, and C. Zhang, "Blind separation of superimposed moving images using image statistics," *TPAMI*, , no. 99, pp. 1–1.

[17] Y. Weiss, "Deriving intrinsic images from image sequences," in *Proceedings of ICCV*. Citeseer, 2001, vol. 2, pp. 68–75.

[18] A. Nemirovski, "Efficient methods in convex programming," 2005.

[19] Y. Nesterov and I.U.E. Nesterov, *Introductory lectures on convex optimization: A basic course*, vol. 87, Springer, 2004.

[20] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[21] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *TIP*, vol. 18, no. 11, pp. 2419–2434, 2009.

[22] H. Li, K Gai, P Gong, and C Zhang, "Effcient superimposition recovering algorithm," *Arxiv.org 1211.4307*, 2012.