Tianchai *Fixed Point Theory Algorithms Sci Eng* (2021) 2021:18
https://doi.org/10.1186/s13663-021-00703-6

Fixed Point Theory and Algorithms for Sciences and Engineering

RESEARCH Open Access

An improved fast iterative shrinkage thresholding algorithm with an error for image deblurring problem

Pattanapong Tianchai[1*]

*Correspondence: pattana@mju.ac.th
[1]Faculty of Science, Maejo University, Chiangmai 50290, Thailand

Abstract

In this paper, we introduce a new iterative forward-backward splitting method with an error for solving the variational inclusion problem of the sum of two monotone operators in real Hilbert spaces. We suggest and analyze this method under some mild appropriate conditions imposed on the parameters such that another strong convergence theorem for these problem is obtained. We also apply our main result to improve the fast iterative shrinkage thresholding algorithm (IFISTA) with an error for solving the image deblurring problem. Finally, we provide numerical experiments to illustrate the convergence behavior and show the effectiveness of the sequence constructed by the inertial technique to the fast processing with high performance and the fast convergence with good performance of IFISTA.

MSC: 47H05; 47H09; 47H20

Keywords: Variational inclusion problem; Maximal monotone operator; Forward-backward method; Iterative shrinkage thresholding; Image deblurring problem

1 Introduction

Let C be a nonempty closed convex subset of a real Hilbert space H. The variational inclusion problem is a fundamental problem in optimization theory, it can be applied in many areas of science and applied science, engineering, economics, and medicine [1–9], in image processing, machine learning, modeling intensity modulated radiation theory treatment planning [10–15]. It is to find $x^* \in H$ such that

$$0 \in Ax^* + Bx^*, \tag{1.1}$$

where $A : H \to H$ is an operator and $B : D(B) \subset H \to 2^H$ is a set-valued operator.

To solve the variational inclusion problem (1.1) via fixed point theory, we define the mapping $J_r^{A,B} : H \to D(B)$ as follows:

$$J_r^{A,B} = (I + rB)^{-1}(I - rA) = J_r^B(I - rA),$$

where $J_r^B = (I + rB)^{-1}$ is the resolvent operator of B for $r > 0$. For $x \in H$, we see that

$$\begin{aligned} J_r^{A,B}(x) = x \quad &\Leftrightarrow \quad x = (I + rB)^{-1}(x - rAx) \\ &\Leftrightarrow \quad x - rAx \in x + rBx \\ &\Leftrightarrow \quad 0 \in Ax + Bx, \end{aligned}$$

which shows that the fixed point set of $J_r^{A,B}$ coincides with the solutions set of $(A + B)^{-1}(0)$. This suggests the following iteration process: $x_1 \in C$ and

$$x_{n+1} = \underbrace{(I + r_n B)^{-1}}_{\text{backward step}} \underbrace{(I - r_n A)}_{\text{forward step}} x_n = J_{r_n}^{A,B}(x_n), \quad \forall n \in \mathbb{N},$$

where $\{r_n\} \subset (0, \infty)$ and $D(B) \subset C$. This method is called a forward-backward splitting algorithm [16, 17].

In applications, we always let $A = \nabla F$ and $B = \partial G$ such that $F : H \to \mathbb{R}$ is a convex and differentiable function and $G : H \to \mathbb{R} \cup \{+\infty\}$ is a convex and lower semi-continuous function, where ∇F is the gradient of F with L-Lipschitz continuous and ∂G is the subdifferential of G which defined by

$$\partial G(x) = \bigl\{z \in H : \langle y - x, z\rangle + G(x) \le G(y), \forall y \in H\bigr\}.$$

Then problem (1.1) is reduced to the following convex minimization problem:

$$F\bigl(x^*\bigr) + G\bigl(x^*\bigr) = \min_{x\in H}\bigl\{F(x) + G(x)\bigr\} \quad \Leftrightarrow \quad 0 \in \nabla F\bigl(x^*\bigr) + \partial G\bigl(x^*\bigr). \tag{1.2}$$

Recall that the proximity operator prox_G of G is defined for all $x \in H$ as follows:

$$\operatorname{prox}_G(x) = \operatorname*{Argmin}_{y\in H}\left\{G(y) + \frac{1}{2}\|y - x\|_2^2\right\}.$$

For $x \in H$ and $r > 0$, we see that

$$\begin{aligned} z = \operatorname{prox}_{rG}(x) \quad &\Leftrightarrow \quad 0 \in r\partial G(z) + z - x \\ &\Leftrightarrow \quad x \in (I + r\partial G)(z) \\ &\Leftrightarrow \quad z = (I + r\partial G)^{-1}(x) = J_r^{\partial G}(x). \end{aligned}$$

Therefore,

$$\begin{aligned} x^* \in \operatorname*{Argmin}_{x\in H}\bigl\{F(x) + G(x)\bigr\} \quad &\Leftrightarrow \quad 0 \in \nabla F\bigl(x^*\bigr) + \partial G\bigl(x^*\bigr) \\ &\Leftrightarrow \quad x^* = J_r^{\nabla F,\partial G}\bigl(x^*\bigr) \\ &\Leftrightarrow \quad x^* = J_r^{\partial G}(I - r\nabla F)x^* = \operatorname{prox}_{rG}(I - r\nabla F)x^*. \end{aligned}$$

Many researchers have proposed and analyzed the iterative shrinkage thresholding algorithms for solving the convex minimization problem (1.2) under a few specific conditions as follows.

In the weak convergence theorems, Lions and Mercier [16] first introduced forward-backward splitting (FBS) algorithm:

$$x_{n+1} = \operatorname{prox}_{\lambda_n G}\bigl(x_n - \lambda_n \nabla F(x_n)\bigr), \quad \forall n \in \mathbb{N},$$

where $x_1 \in H$ and $\{\lambda_n\} \subset (0, 2/L)$. Later, Moudafi and Oliny [18] introduced the iterative forward-backward splitting (IFBS) algorithm:

$$\begin{cases} y_n = x_n + \theta_n(x_n - x_{n-1}), \\ x_{n+1} = \operatorname{prox}_{\lambda_n G}(y_n - \lambda_n \nabla F(x_n)), \quad \forall n \in \mathbb{N}, \end{cases}$$

where $x_0, x_1 \in H$, $\{\theta_n\} \subset [0,a] \subset [0,1)$, $\{\lambda_n\} \subset [b,c] \subset (0,2/L)$ such that $\sum_{n=1}^{\infty} \theta_n \|x_n - x_{n-1}\|^2 < \infty$. In our research, we focus attention on the inertial parameter θ_n which controls the momentum of $x_n - x_{n-1}$ in the fast iterative shrinkage thresholding algorithm (FISTA) of Beck and Teboulle [19] as follows:

$$\begin{cases} x_n = \operatorname{prox}_{\frac{1}{L}G}(y_n - \frac{1}{L}\nabla F(y_n)), \\ t_{n+1} = \frac{1+\sqrt{1+4t_n^2}}{2}, \qquad \theta_n = \frac{t_n - 1}{t_{n+1}}, \\ y_{n+1} = x_n + \theta_n(x_n - x_{n-1}), \quad \forall n \in \mathbb{N}, \end{cases}$$

where $y_1 = x_0 \in H$ and $t_1 = 1$. In FISTA, we observe that y_n is known before x_n, where the sequence $\{x_n\}$ converges weakly to the solution of the convex minimization problem (1.2). Recently, Hanjing and Suantai [20] introduced the forward-backward modified W-algorithm (FBMWA) as follows:

$$\begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}), \\ z_n = (1-\gamma_n)w_n + \gamma_n \operatorname{prox}_{\lambda_n G}(w_n - \lambda_n \nabla F(w_n)), \\ y_n = (1-\beta_n)\operatorname{prox}_{\lambda_n G}(w_n - \lambda_n \nabla F(w_n)) + \beta_n \operatorname{prox}_{\lambda_n G}(z_n - \lambda_n \nabla F(z_n)), \\ x_{n+1} = (1-\alpha_n)\operatorname{prox}_{\lambda_n G}(z_n - \lambda_n \nabla F(z_n)) + \alpha_n \operatorname{prox}_{\lambda_n G}(y_n - \lambda_n \nabla F(y_n)), \end{cases}$$

for all $n \in \mathbb{N}$, where $x_0, x_1 \in H$ and $\{\alpha_n\} \subset [0,a] \subset [0,1)$, $\{\beta_n\} \subset [0,1]$, $\{\gamma_n\} \subset [b,c] \subset (0,1)$, and $\{\theta_n\} \subset [0,\infty)$ such that $\sum_{n=1}^{\infty} \theta_n < \infty$, and $\{\lambda_n\} \subset (0,2/L)$ such that $\lambda_n \to \lambda \in (0,2/L)$ as $n \to \infty$. In the same way, Padcharoen and Kuman [21] introduced the forward-backward modified MM-algorithm (FBMMMA) as follows:

$$\begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}), \\ z_n = (1-\gamma_n)w_n + \gamma_n \operatorname{prox}_{\lambda_n G}(w_n - \lambda_n \nabla F(w_n)), \\ y_n = (1-\alpha_n-\beta_n)z_n + \alpha_n \operatorname{prox}_{\lambda_n G}(z_n - \lambda_n \nabla F(z_n)) \\ \qquad + \beta_n \operatorname{prox}_{\lambda_n G}(w_n - \lambda_n \nabla F(w_n)), \\ x_{n+1} = \operatorname{prox}_{\lambda_n G}(y_n - \lambda_n \nabla F(y_n)), \quad \forall n \in \mathbb{N}, \end{cases}$$

where $x_0, x_1 \in H$ and $\{\alpha_n\}, \{\beta_n\}, \{\gamma_n\} \subset [0,1]$ such that $\alpha_n + \beta_n \in [0,1]$ and $\{\theta_n\} \subset (0,1)$ such that $\sum_{n=1}^{\infty} \theta_n < \infty$, and $\{\lambda_n\} \subset (0,2/L)$ such that $\lambda_n \to \lambda \in (0,2/L)$ as $n \to \infty$. Other weak convergence theorems of all those algorithms are obtained.

In the strong convergence theorems, Verma and Shukla [22] introduced a new accelerated proximal gradient algorithm (NAGA) as follows:

$$\begin{cases} z_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = (1-\alpha_n)z_n + \alpha_n \operatorname{prox}_{\lambda_n G}(z_n - \lambda_n \nabla F(z_n)), \\ x_{n+1} = \operatorname{prox}_{\lambda_n G}(y_n - \lambda_n \nabla F(y_n)), \quad \forall n \in \mathbb{N}, \end{cases}$$

where $x_0, x_1 \in H$, $\{\alpha_n\}, \{\theta_n\} \subset (0,1]$, and $\{\lambda_n\} \subset (0, 2/L)$. They proved that the sequence $\{x_n\}$ of NAGA converges strongly under the condition $\frac{\|x_n - x_{n-1}\|}{\theta_n} \to 0$ as $n \to \infty$. How to choose the parameter θ_n? We leave it for the reader to verify. In their proof, we observe that NAGA still holds under conditions $\alpha_n \to 0$ and $\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| \to 0$ as $n \to \infty$, and the parameter θ_n can be chosen as

$$\theta_n = \begin{cases} \min\{\frac{\omega_n}{\|x_n - x_{n-1}\|}, \alpha_n\} & \text{if } x_n \neq x_{n-1}, \\ \alpha_n & \text{otherwise}, \end{cases}$$

where $\{\omega_n\}$ is a positive sequence such that $\omega_n = o(\alpha_n)$. Cholamjiak et al. [23] introduced the strong convergence theorem for the inclusion problem (SCTIP) by letting $S = I$, $A = \nabla F$, and $B = \partial G$ as follows:

$$\begin{cases} z_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = \alpha_n f(x_n) + (1-\alpha_n) \operatorname{prox}_{\lambda_n G}(z_n - \lambda_n \nabla F(z_n)), \\ x_{n+1} = \beta_n x_n + (1-\beta_n) y_n, \quad \forall n \in \mathbb{N}, \end{cases}$$

where $x_0, x_1 \in C$ and f is a contraction of C into itself, and $\{\alpha_n\}, \{\beta_n\} \subset (0,1)$, $\{\lambda_n\} \subset (0, 2/L)$, and $\{\theta_n\} \subset [0, \theta]$ such that $\theta \in [0,1)$. They proved that the sequence $\{x_n\}$ of SCTIP converges strongly under the following conditions:

(C1) $\lim_{n\to\infty} \alpha_n = 0$ and $\sum_{n=1}^{\infty} \alpha_n = \infty$,
(C2) $\liminf_{n\to\infty} \beta_n(1-\beta_n) > 0$,
(C3) $0 < \liminf_{n\to\infty} \lambda_n \leq \limsup_{n\to\infty} \lambda_n < 2/L$,
(C4) $\lim_{n\to\infty} \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| = 0$.

Moreover, many researchers have proposed and analyzed the iterative forward-backward scheme with a variable step size, which does not depend on the Lipschitz constant of the operator $A = \nabla F$ (see also [24, 25]).

In our research, we consider the forward-backward splitting method with an error as follows: $x_1 \in C$ and

$$x_{n+1} = \underbrace{(I + \lambda_n B)^{-1}}_{\text{backward step}} \underbrace{\big((I - \lambda_n A)x_n + e_n\big)}_{\text{forward step with an error}} = J_{\lambda_n}^B \big((I - \lambda_n A)x_n + e_n\big), \quad \forall n \in \mathbb{N},$$

where $\{\lambda_n\} \subset (0, \infty)$, $\{e_n\} \subset H$, $D(B) \subset C$, and $J_{\lambda_n}^B = (I + \lambda_n B)^{-1}$. We introduce a new iterative forward-backward splitting method with an error for solving the variational inclusion problem (1.1) as follows:

$$\begin{cases} z_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = \alpha_n f(z_n) + (1-\alpha_n)J^B_{\lambda_n}(z_n - \lambda_n A z_n + e_n), \\ x_{n+1} = J^B_{\lambda_n}(y_n - \lambda_n A y_n + e_n), \quad \forall n \in \mathbb{N}, \end{cases}$$

where $x_0, x_1 \in C$ and f is a contraction of C into itself, and $\{\alpha_n\} \subset (0,1)$, $\{\lambda_n\} \subset (0, 2/L)$, $\{e_n\} \subset H$, and $\{\theta_n\} \subset [0,\theta]$ such that $\theta \in [0,1)$. Moreover, it can be applied to improve the fast iterative shrinkage thresholding algorithm (IFISTA) with an error for solving the convex minimization problem (1.2) by letting $A = \nabla F$ and $B = \partial G$ as follows:

$$\begin{cases} z_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = \alpha_n f(z_n) + (1-\alpha_n)\operatorname{prox}_{\lambda_n G}(z_n - \lambda_n \nabla F(z_n) + e_n), \\ x_{n+1} = \operatorname{prox}_{\lambda_n G}(y_n - \lambda_n \nabla F(y_n) + e_n), \quad \forall n \in \mathbb{N} \end{cases}$$

which obtains a self-adaptive scheme with fast convergence properties under some mild conditions when compared to the existing algorithms in the literature. The outline of our research is as follows: in Sect. 2, we give some well-known definitions and lemmas which are used in Sect. 3 to prove the strong convergence theorem of IFISTA for solving the variational inclusion problem (1.1), and we also apply its result in Sect. 4 for solving the image deblurring problem, which is a special case of convex minimization problem (1.2); and in Sect. 5, we provide numerical experiments to illustrate the fast processing with high performance and the fast convergence with good performance of IFISTA by the inertial technique.

2 Preliminaries

Let C be a nonempty closed convex subset of a real Hilbert space H. We will use the notation: $\to$ to denote the strong convergence, $\rightharpoonup$ to denote the weak convergence,

$$\omega_w(x_n) = \bigl\{x : \exists\{x_{n_k}\} \subset \{x_n\} \text{ such that } x_{n_k} \rightharpoonup x\bigr\}$$

to denote the weak limit set of $\{x_n\}$, and $\operatorname{Fix}(T) = \{x : x = Tx\}$ to denote the fixed point set of the mapping T.

Recall that the metric projection $P_C : H \to C$ is defined as follows: for each $x \in H$, $P_C x$ is the unique point in C satisfying

$$\|x - P_C x\| = \inf\bigl\{\|x - y\| : y \in C\bigr\}.$$

The operator $T : H \to H$ is called:

(i) monotone if

$$\langle x - y, Tx - Ty\rangle \geq 0, \quad \forall x, y \in H,$$

(ii) L-Lipschitzian with $L > 0$ if

$$\|Tx - Ty\| \leq L\|x - y\|, \quad \forall x, y \in H,$$

(iii) k-contraction if it is k-Lipschitzian with $k \in (0,1)$,
(iv) nonexpansive if it is 1-Lipschitzian,
(v) firmly nonexpansive if

$$\|Tx - Ty\|^2 \leq \|x - y\|^2 - \left\|(I - T)x - (I - T)y\right\|^2, \quad \forall x, y \in H,$$

(vi) α-strongly monotone with $\alpha > 0$ if

$$\langle Tx - Ty, x - y\rangle \geq \alpha \|x - y\|^2, \quad \forall x, y \in H,$$

(vii) α-inverse strongly monotone with $\alpha > 0$ if

$$\langle Tx - Ty, x - y\rangle \geq \alpha \|Tx - Ty\|^2, \quad \forall x, y \in H.$$

Let B be a mapping of H into 2^H. The domain and the range of B are denoted by $D(B) = \{x \in H : Bx \neq \emptyset\}$ and $R(B) = \bigcup\{Bx : x \in D(B)\}$, respectively. The inverse of B, denoted by B^{-1}, is defined by $x \in B^{-1}y$ if and only if $y \in Bx$. A multi-valued mapping B is said to be a monotone operator on H if $\langle x - y, u - v\rangle \geq 0$ for all $x, y \in D(B)$, $u \in Bx$, and $v \in By$. A monotone operator B on H is said to be maximal if its graph is not strictly contained in the graph of any other monotone operator on H. For a maximal monotone operator B on H and $r > 0$, we define the single-valued resolvent operator $J_r^B : H \to D(B)$ by $J_r^B = (I + rB)^{-1}$. It is well known that J_r^B is firmly nonexpansive and $\mathrm{Fix}(J_r^B) = B^{-1}(0)$.

We collect together some known lemmas which are the main tools in proving our result.

Lemma 2.1 ([26]) *Let C be a nonempty closed convex subset of a real Hilbert space H. Then*:
(i) $\|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y\rangle + \|y\|^2$, $\forall x, y \in H$,
(ii) $\|\lambda x + (1 - \lambda)y\|^2 = \lambda\|x\|^2 + (1 - \lambda)\|y\|^2 - \lambda(1 - \lambda)\|x - y\|^2$, $\forall x, y \in H$, $\lambda \in \mathbb{R}$,
(iii) $z = P_C x \Leftrightarrow \langle x - z, z - y\rangle \geq 0$, $\forall x \in H$, $y \in C$,
(iv) $z = P_C x \Leftrightarrow \|x - z\|^2 \leq \|x - y\|^2 - \|y - z\|^2$, $\forall x \in H$, $y \in C$,
(v) $\|P_C x - P_C y\|^2 \leq \langle x - y, P_C x - P_C y\rangle$, $\forall x, y \in H$.

Lemma 2.2 ([27]) *Let H and K be two real Hilbert spaces, and let $T : K \to K$ be a firmly nonexpansive mapping such that $\|(I - T)x\|$ is a convex function from K to $\overline{\mathbb{R}} = [-\infty, +\infty]$. Let $A : H \to K$ be a bounded linear operator and $f(x) = \frac{1}{2}\|(I - T)Ax\|^2$ for all $x \in H$. Then*:
(i) *f is convex and differential*,
(ii) *$\nabla f(x) = A^*(I - T)Ax$ for all $x \in H$ such that A^* denotes the adjoint of A*,
(iii) *f is weakly lower semi-continuous on H*,
(iv) *∇f is $\|A\|^2$-Lipschitzian.*

Lemma 2.3 ([27]) *Let H be a real Hilbert space and $T : H \to H$ be an operator. The following statements are equivalent*:
(i) *T is firmly nonexpansive*,
(ii) *$\|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty\rangle$, $\forall x, y \in H$*,
(iii) *$I - T$ is firmly nonexpansive.*

Lemma 2.4 ([28]) *Let C be a nonempty closed convex subset of a real Hilbert space H. Let the mapping $A : C \to H$ be α-inverse strongly monotone and $r > 0$ be a constant. Then we have*

$$\left\|(I - rA)x - (I - rA)y\right\|^2 \le \|x - y\|^2 - r(2\alpha - r)\|Ax - Ay\|^2$$

for all $x, y \in C$. In particular, if $0 < r \le 2\alpha$, then $I - rA$ is nonexpansive.

Lemma 2.5 ([29] (Demiclosedness principle)) *Let C be a nonempty closed convex subset of a real Hilbert space H, and let $S : C \to C$ be a nonexpansive mapping with $Fix(S) \neq \emptyset$. If the sequence $\{x_n\} \subset C$ converges weakly to x and the sequence $\{(I - S)x_n\}$ converges strongly to y, then $(I - S)x = y$; in particular, if $y = 0$, then $x \in Fix(S)$.*

Lemma 2.6 ([30]) *Let $\{a_n\}$ and $\{c_n\}$ be sequences of nonnegative real numbers such that*

$$a_{n+1} \le (1 - \delta_n)a_n + b_n + c_n, \quad \forall n = 0, 1, 2, \dots,$$

where $\{\delta_n\}$ is a sequence in $(0, 1)$ and $\{b_n\}$ is a real sequence. Assume that $\sum_{n=0}^{\infty} c_n < \infty$. Then the following results hold:

(i) *if $b_n \le \delta_n M$ for some $M \ge 0$, then $\{a_n\}$ is a bounded sequence,*
(ii) *if $\sum_{n=0}^{\infty} \delta_n = \infty$ and $\limsup_{n\to\infty} b_n/\delta_n \le 0$, then $\lim_{n\to\infty} a_n = 0$.*

Lemma 2.7 ([31]) *Assume that $\{s_n\}$ is a sequence of nonnegative real numbers such that*

$$s_{n+1} \le (1 - \gamma_n)s_n + \gamma_n\delta_n, \quad \forall n = 0, 1, 2, \dots$$

and

$$s_{n+1} \le s_n - \eta_n + \rho_n, \quad \forall n = 0, 1, 2, \dots,$$

where $\{\gamma_n\}$ is a sequence in $(0, 1)$, $\{\eta_n\}$ is a sequence of nonnegative real numbers, and $\{\delta_n\}$, $\{\rho_n\}$ are real sequences such that

(i) $\sum_{n=0}^{\infty} \gamma_n = \infty$,
(ii) $\lim_{n\to\infty} \rho_n = 0$,
(iii) *if $\lim_{k\to\infty} \eta_{n_k} = 0$, then $\limsup_{k\to\infty} \delta_{n_k} \le 0$ for any subsequence $\{n_k\}$ of $\{n\}$.*

Then $\lim_{n\to\infty} s_n = 0$.

3 Main result

Theorem 3.1 *Let C be a nonempty closed convex subset of a real Hilbert space H. Let A be an α-inverse strongly monotone mapping of H into itself and B be a maximal monotone operator on H such that the domain of B is included in C, and assume that $(A + B)^{-1}(0)$ is nonempty. Let $J_\lambda^B = (I + \lambda B)^{-1}$ be the resolvent of B for $\lambda > 0$ and f be a k-contraction mapping of C into itself. Let $x_0, x_1 \in C$ and $\{x_n\} \subset C$ be a sequence generated by*

$$\begin{cases} z_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = \alpha_n f(z_n) + (1 - \alpha_n)J_{\lambda_n}^B(z_n - \lambda_n Az_n + e_n), \\ x_{n+1} = J_{\lambda_n}^B(y_n - \lambda_n Ay_n + e_n), \end{cases}$$

for all $n \in \mathbb{N}$, *where* $\{\alpha_n\} \subset (0,1)$, $\{\lambda_n\} \subset (0,2\alpha)$, $\{e_n\} \subset H$, *and* $\{\theta_n\} \subset [0,\theta]$ *such that* $\theta \in [0,1)$ *satisfy the following conditions*:

(C1) $\lim_{n\to\infty} \alpha_n = 0$ *and* $\sum_{n=1}^{\infty} \alpha_n = \infty$,

(C2) $0 < a \leq \lambda_n \leq b < 2\alpha$ *for some* $a,b > 0$,

(C3) $\lim_{n\to\infty} \frac{\|e_n\|}{\alpha_n} = 0$,

(C4) $\sum_{n=1}^{\infty} \|e_n\| < \infty$ *and* $\lim_{n\to\infty} \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| = 0$.

Then the sequence $\{x_n\}$ *converges strongly to a point* $x^* \in (A+B)^{-1}(0)$ *where* $x^* = P_{(A+B)^{-1}(0)}f(x^*)$.

Proof Picking $z \in (A+B)^{-1}(0)$ and fixing $n \in \mathbb{N}$, it follows that $z = J_{\lambda_n}^B(z - \lambda_n Az)$. Firstly, we will show that $\{x_n\}$, $\{y_n\}$, and $\{z_n\}$ are bounded. Since

$$\|z_n - z\| \leq \|x_n - z\| + \theta_n\|x_n - x_{n-1}\|,$$

therefore, by nonexpansiveness of $J_{\lambda_n}^B$ and $I - \lambda_n A$, we have

$$\begin{aligned}
\|y_n - z\| &= \left\|\alpha_n\big(f(z_n) - z\big) + (1-\alpha_n)\big(J_{\lambda_n}^B(z_n - \lambda_n Az_n + e_n) - z\big)\right\| \\
&\leq \alpha_n\big(\left\|f(z_n) - f(z)\right\| + \left\|f(z) - z\right\|\big) \\
&\quad + (1-\alpha_n)\left\|(z_n - \lambda_n Az_n + e_n) - (z - \lambda_n Az)\right\| \\
&\leq \alpha_n\big(k\|z_n - z\| + \left\|f(z) - z\right\|\big) + (1-\alpha_n)\big(\|z_n - z\| + \|e_n\|\big) \\
&\leq \big(1 - \alpha_n(1-k)\big)\|z_n - z\| + \alpha_n\left\|f(z) - z\right\| + \|e_n\| \\
&\leq \big(1 - \alpha_n(1-k)\big)\|x_n - z\| + \theta_n\|x_n - x_{n-1}\| + \alpha_n\left\|f(z) - z\right\| + \|e_n\|.
\end{aligned}$$

It follows by the same arguments again that

$$\begin{aligned}
\|x_{n+1} - z\| &= \left\|J_{\lambda_n}^B(y_n - \lambda_n Ay_n + e_n) - J_{\lambda_n}^B(z - \lambda_n Az)\right\| \\
&\leq \left\|(y_n - \lambda_n Ay_n + e_n) - (z - \lambda_n Az)\right\| \\
&\leq \|y_n - z\| + \|e_n\| \\
&\leq \big(1 - \alpha_n(1-k)\big)\|x_n - z\| \\
&\quad + \alpha_n(1-k)\left(\frac{1}{1-k}\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| + \frac{\|f(z) - z\|}{1-k}\right) + 2\|e_n\|.
\end{aligned}$$

So, by condition (C4) and putting $M = \frac{1}{1-k}(\|f(z) - z\| + \sup_{n\in\mathbb{N}} \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\|) \geq 0$ in Lemma 2.6 (i), we conclude that the sequence $\{\|x_n - z\|\}$ is bounded. That is the sequence $\{x_n\}$ is bounded, and so is $\{z_n\}$. Moreover, by condition (C4), $\sum_{n=1}^{\infty} \|e_n\| < \infty$ implies $\lim_{n\to\infty} \|e_n\| = 0$, that is, $\lim_{n\to\infty} e_n = 0$, it follows that the sequence $\{y_n\}$ is also bounded.

Since $P_{(A+B)^{-1}(0)}f$ is k-contraction on C, by Banach's contraction principle there exists a unique element $x^* \in C$ such that $x^* = P_{(A+B)^{-1}(0)}f(x^*)$, that is, $x^* \in (A+B)^{-1}(0)$, it follows that $x^* = J_{\lambda_n}^B(x^* - \lambda_n Ax^*)$. Now, we will show that $x_n \to x^*$ as $n \to \infty$. On the other hand, we have

$$\begin{aligned}
\left\|z_n - x^*\right\|^2 &= \left\langle z_n - x^*, z_n - x^*\right\rangle \\
&= \left\langle x_n + \theta_n(x_n - x_{n-1}) - x^*, z_n - x^*\right\rangle
\end{aligned}$$

$$
\begin{aligned}
&= \langle x_n - x^*, z_n - x^* \rangle + \theta_n \langle x_n - x_{n-1}, z_n - x^* \rangle \\
&\leq \|x_n - x^*\| \|z_n - x^*\| + \theta_n \|x_n - x_{n-1}\| \|z_n - x^*\| \\
&\leq \frac{1}{2}\left(\|x_n - x^*\|^2 + \|z_n - x^*\|^2\right) + \theta_n \|x_n - x_{n-1}\| \|z_n - x^*\|.
\end{aligned}
$$

This implies that

$$
\|z_n - x^*\|^2 \leq \|x_n - x^*\|^2 + 2\theta_n \|x_n - x_{n-1}\| \|z_n - x^*\|. \tag{3.1}
$$

It follows by (3.1), Lemma 2.4, and the firm nonexpansiveness of $J^B_{\lambda_n}$ that

$$
\begin{aligned}
&\left\|J^B_{\lambda_n}(z_n - \lambda_n A z_n + e_n) - x^*\right\|^2 \\
&\quad = \left\|J^B_{\lambda_n}(z_n - \lambda_n A z_n + e_n) - J^B_{\lambda_n}\left(x^* - \lambda_n A x^*\right)\right\|^2 \\
&\quad \leq \left\|(z_n - \lambda_n A z_n + e_n) - \left(x^* - \lambda_n A x^*\right)\right\|^2 \\
&\qquad - \left\|\left(I - J^B_{\lambda_n}\right)(z_n - \lambda_n A z_n + e_n) - \left(I - J^B_{\lambda_n}\right)\left(x^* - \lambda_n A x^*\right)\right\|^2 \\
&\quad \leq \left(\left\|(z_n - \lambda_n A z_n) - \left(x^* - \lambda_n A x^*\right)\right\| + \|e_n\|\right)^2 \\
&\qquad - \left\|\left(I - J^B_{\lambda_n}\right)(z_n - \lambda_n A z_n + e_n) - \left(I - J^B_{\lambda_n}\right)\left(x^* - \lambda_n A x^*\right)\right\|^2 \\
&\quad = \left\|(I - \lambda_n A) z_n - (I - \lambda_n A) x^*\right\|^2 + 2\left\|(z_n - \lambda_n A z_n) - \left(x^* - \lambda_n A x^*\right)\right\| \|e_n\| \\
&\qquad + \|e_n\|^2 - \left\|\left(I - J^B_{\lambda_n}\right)(z_n - \lambda_n A z_n + e_n) - \left(I - J^B_{\lambda_n}\right)\left(x^* - \lambda_n A x^*\right)\right\|^2 \\
&\quad \leq \|z_n - x^*\|^2 - \lambda_n(2\alpha - \lambda_n)\|A z_n - A x^*\|^2 \\
&\qquad + 2\left\|(z_n - \lambda_n A z_n) - \left(x^* - \lambda_n A x^*\right)\right\| \|e_n\| + \|e_n\|^2 \\
&\qquad - \left\|\left(I - J^B_{\lambda_n}\right)(z_n - \lambda_n A z_n + e_n) - \left(I - J^B_{\lambda_n}\right)\left(x^* - \lambda_n A x^*\right)\right\|^2 \\
&\quad \leq \|x_n - x^*\|^2 + 2\theta_n \|x_n - x_{n-1}\| \|z_n - x^*\| - \lambda_n(2\alpha - \lambda_n)\|A z_n - A x^*\|^2 \\
&\qquad + 2\left\|(z_n - \lambda_n A z_n) - \left(x^* - \lambda_n A x^*\right)\right\| \|e_n\| + \|e_n\|^2 \\
&\qquad - \left\|\left(I - J^B_{\lambda_n}\right)(z_n - \lambda_n A z_n + e_n) - \left(I - J^B_{\lambda_n}\right)\left(x^* - \lambda_n A x^*\right)\right\|^2.
\end{aligned} \tag{3.2}
$$

We also have

$$
\begin{aligned}
\|y_n - x^*\|^2 &= \langle y_n - x^*, y_n - x^* \rangle \\
&= \left\langle \alpha_n f(z_n) + (1 - \alpha_n) J^B_{\lambda_n}(z_n - \lambda_n A z_n + e_n) - x^*, y_n - x^* \right\rangle \\
&= \left\langle \alpha_n \left(f(z_n) - x^*\right) + (1 - \alpha_n)\left(J^B_{\lambda_n}(z_n - \lambda_n A z_n + e_n) - x^*\right), y_n - x^* \right\rangle \\
&= \alpha_n \left\langle f(z_n) - f\left(x^*\right), y_n - x^* \right\rangle + \alpha_n \left\langle f\left(x^*\right) - x^*, y_n - x^* \right\rangle \\
&\quad + (1 - \alpha_n)\left\langle J^B_{\lambda_n}(z_n - \lambda_n A z_n + e_n) - x^*, y_n - x^* \right\rangle \\
&\leq \alpha_n k \|z_n - x^*\| \|y_n - x^*\| + \alpha_n \left\langle f\left(x^*\right) - x^*, y_n - x^* \right\rangle \\
&\quad + (1 - \alpha_n)\left\|J^B_{\lambda_n}(z_n - \lambda_n A z_n + e_n) - x^*\right\| \|y_n - x^*\| \\
&\leq \frac{1}{2}\alpha_n k\left(\|z_n - x^*\|^2 + \|y_n - x^*\|^2\right) + \alpha_n \left\langle f\left(x^*\right) - x^*, y_n - x^* \right\rangle
\end{aligned}
$$

$$+\frac{1}{2}(1-\alpha_n)\left(\left\|J_{\lambda_n}^B(z_n-\lambda_n Az_n+e_n)-x^*\right\|^2+\left\|y_n-x^*\right\|^2\right).$$

This implies that

$$\begin{aligned}\left\|y_n-x^*\right\|^2 &\le \frac{\alpha_n k}{1+\alpha_n(1-k)}\left\|z_n-x^*\right\|^2+\frac{2\alpha_n}{1+\alpha_n(1-k)}\left\langle f\left(x^*\right)-x^*,y_n-x^*\right\rangle\\ &\quad+\frac{1-\alpha_n}{1+\alpha_n(1-k)}\left\|J_{\lambda_n}^B(z_n-\lambda_n Az_n+e_n)-x^*\right\|^2. \end{aligned}\tag{3.3}$$

Hence, by (3.1), (3.2), (3.3), the nonexpansiveness of $J_{\lambda_n}^B$, and $I-\lambda_n A$, we obtain

$$\begin{aligned}\left\|x_{n+1}-x^*\right\|^2 &= \left\|J_{\lambda_n}^B(y_n-\lambda_n Ay_n+e_n)-J_{\lambda_n}^B\left(x^*-\lambda_n Ax^*\right)\right\|^2\\ &\le \left\|(y_n-\lambda_n Ay_n+e_n)-\left(x^*-\lambda_n Ax^*\right)\right\|^2\\ &\le \left(\left\|y_n-x^*\right\|+\|e_n\|\right)^2\\ &= \left\|y_n-x^*\right\|^2+2\left\|y_n-x^*\right\|\|e_n\|+\|e_n\|^2\\ &\le \frac{\alpha_n k}{1+\alpha_n(1-k)}\left(\left\|x_n-x^*\right\|^2+2\theta_n\|x_n-x_{n-1}\|\left\|z_n-x^*\right\|\right)\\ &\quad+2\left\|y_n-x^*\right\|\|e_n\|+\|e_n\|^2+\frac{2\alpha_n}{1+\alpha_n(1-k)}\left\langle f\left(x^*\right)-x^*,y_n-x^*\right\rangle\\ &\quad+\frac{1-\alpha_n}{1+\alpha_n(1-k)}\left(\left\|x_n-x^*\right\|^2+2\theta_n\|x_n-x_{n-1}\|\left\|z_n-x^*\right\|\right.\\ &\quad-\lambda_n(2\alpha-\lambda_n)\left\|Az_n-Ax^*\right\|^2+2\left\|(z_n-\lambda_n Az_n)-\left(x^*-\lambda_n Ax^*\right)\right\|\|e_n\|\\ &\quad\left.+\|e_n\|^2-\left\|\left(I-J_{\lambda_n}^B\right)(z_n-\lambda_n Az_n+e_n)-\left(I-J_{\lambda_n}^B\right)\left(x^*-\lambda_n Ax^*\right)\right\|^2\right). \end{aligned}$$

It follows that

$$\begin{aligned}\left\|x_{n+1}-x^*\right\|^2 &\le \left(1-\frac{\alpha_n(1-k)}{1+\alpha_n(1-k)}\right)\left\|x_n-x^*\right\|^2\\ &\quad+\frac{\alpha_n(1-k)}{1+\alpha_n(1-k)}\left(\frac{2}{1-k}\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\|\left\|z_n-x^*\right\|\right.\\ &\quad+\frac{4}{1-k}\frac{\|e_n\|}{\alpha_n}\left\|y_n-x^*\right\|+\frac{2}{1-k}\frac{\|e_n\|}{\alpha_n}\|e_n\|+\frac{2}{1-k}\left\langle f\left(x^*\right)-x^*,y_n-x^*\right\rangle\\ &\quad\left.+\frac{2}{1-k}\frac{\|e_n\|}{\alpha_n}\left\|(z_n-\lambda_n Az_n)-\left(x^*-\lambda_n Ax^*\right)\right\|\right) \end{aligned}$$

and

$$\begin{aligned}\left\|x_{n+1}-x^*\right\|^2 &\le \left\|x_n-x^*\right\|^2-\left(\lambda_n(2\alpha-\lambda_n)\left\|Az_n-Ax^*\right\|^2\right.\\ &\quad\left.+\left\|\left(I-J_{\lambda_n}^B\right)(z_n-\lambda_n Az_n+e_n)-\left(I-J_{\lambda_n}^B\right)\left(x^*-\lambda_n Ax^*\right)\right\|^2\right)\\ &\quad+\left(2\alpha_n\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\|\left\|z_n-x^*\right\|+2\alpha_n\frac{\|e_n\|}{\alpha_n}\left\|y_n-x^*\right\|+2\|e_n\|^2\right.\\ &\quad\left.+2\alpha_n\left\|f\left(x^*\right)-x^*\right\|\left\|y_n-x^*\right\|+2\left\|(z_n-\lambda_n Az_n)-\left(x^*-\lambda_n Ax^*\right)\right\|\|e_n\|\right), \end{aligned}$$

which are of the forms

$$s_{n+1} \leq (1-\gamma_n)s_n + \gamma_n\delta_n$$

and

$$s_{n+1} \leq s_n - \eta_n + \rho_n,$$

respectively, where $s_n = \|x_n - x^*\|^2$, $\gamma_n = \frac{\alpha_n(1-k)}{1+\alpha_n(1-k)}$, $\delta_n = \frac{2}{1-k}\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\|\|z_n - x^*\| + \frac{4}{1-k}\frac{\|e_n\|}{\alpha_n}\|y_n - x^*\| + \frac{2}{1-k}\frac{\|e_n\|}{\alpha_n}\|e_n\| + \frac{2}{1-k}\langle f(x^*) - x^*, y_n - x^*\rangle + \frac{2}{1-k}\frac{\|e_n\|}{\alpha_n}\|(z_n - \lambda_n Az_n) - (x^* - \lambda_n Ax^*)\|$, $\eta_n = \lambda_n(2\alpha - \lambda_n)\|Az_n - Ax^*\|^2 + \|(I - J^B_{\lambda_n})(z_n - \lambda_n Az_n + e_n) - (I - J^B_{\lambda_n})(x^* - \lambda_n Ax^*)\|^2$ and $\rho_n = 2\alpha_n\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\|\|z_n - x^*\| + 2\alpha_n\frac{\|e_n\|}{\alpha_n}\|y_n - x^*\| + 2\|e_n\|^2 + 2\alpha_n\|f(x^*) - x^*\|\|y_n - x^*\| + 2\|(z_n - \lambda_n Az_n) - (x^* - \lambda_n Ax^*)\|\|e_n\|$. Therefore, using conditions (C1), (C3), and (C4), we can check that all those sequences satisfy conditions (i) and (ii) in Lemma 2.7. To complete the proof, we verify that condition (iii) in Lemma 2.7 is satisfied. Let $\lim_{i\to\infty}\eta_{n_i} = 0$. Then, by condition (C2), we have

$$\lim_{i\to\infty}\left\|Az_{n_i} - Ax^*\right\| = 0 \tag{3.4}$$

and

$$\lim_{i\to\infty}\left\|\left(I - J^B_{\lambda_{n_i}}\right)(z_{n_i} - \lambda_{n_i}Az_{n_i} + e_{n_i}) - \left(I - J^B_{\lambda_{n_i}}\right)\left(x^* - \lambda_{n_i}Ax^*\right)\right\| = 0.$$

It follows by conditions (C2), (C4) and (3.4) that

$$\begin{aligned}
&\lim_{i\to\infty}\Big\|(z_{n_i} - \lambda_{n_i}Az_{n_i} + e_{n_i}) - J^B_{\lambda_{n_i}}(z_{n_i} - \lambda_{n_i}Az_{n_i} + e_{n_i}) \\
&\quad - \left(\left(x^* - \lambda_{n_i}Ax^*\right) - J^B_{\lambda_{n_i}}\left(x^* - \lambda_{n_i}Ax^*\right)\right)\Big\| = 0, \\
&\lim_{i\to\infty}\left\|z_{n_i} - J^B_{\lambda_{n_i}}(z_{n_i} - \lambda_{n_i}Az_{n_i})\right\| = 0.
\end{aligned} \tag{3.5}$$

Consider a subsequence $\{x_{n_i}\}$ of $\{x_n\}$. As $\{x_n\}$ is bounded, so is $\{x_{n_i}\}$, there exists a subsequence $\{x_{n_{i_j}}\}$ of $\{x_{n_i}\}$ which converges weakly to $x \in C$. Without loss of generality, we can assume that $x_{n_i} \rightharpoonup x$ as $i \to \infty$. On the other hand, by conditions (C1) and (C4), we have

$$\lim_{i\to\infty}\|z_{n_i} - x_{n_i}\| = \lim_{i\to\infty}\alpha_{n_i}\frac{\theta_{n_i}}{\alpha_{n_i}}\|x_{n_i} - x_{n_i-1}\| = 0.$$

It follows that $z_{n_i} \rightharpoonup x$ as $i \to \infty$. Hence, by (3.5) and the demiclosedness at zero in Lemma 2.5, we obtain $x \in \mathrm{Fix}(J^B_{\lambda_{n_i}}(I - \lambda_{n_i}A))$, that is, $x \in (A+B)^{-1}(0)$. Since

$$\|y_{n_i} - z_{n_i}\| \leq \alpha_{n_i}\left\|f(z_{n_i}) - z_{n_i}\right\| + (1-\alpha_{n_i})\left\|J^B_{\lambda_{n_i}}(z_{n_i} - \lambda_{n_i}Az_{n_i} + e_{n_i}) - z_{n_i}\right\|,$$

then, by (3.5) and conditions (C1) and (C4), we obtain

$$\lim_{i\to\infty}\|y_{n_i} - z_{n_i}\| = 0.$$

It implies that $y_{n_i} \rightharpoonup x$ as $i \to \infty$. Therefore, by Lemma 2.1(iii), we obtain

$$\limsup_{i\to\infty}\langle f(x^*) - x^*, y_{n_i} - x^*\rangle = \langle f(x^*) - x^*, x - x^*\rangle \le 0.$$

It follows by conditions (C1), (C3), and (C4) that $\limsup_{i\to\infty} \delta_{n_i} \le 0$. So, by Lemma 2.7, we conclude that $x_n \to x^*$ as $n \to \infty$. This completes the proof. □

Remark 3.2 Indeed, the parameter θ_n can be chosen as follows:

$$\theta_n = \begin{cases} \min\{\frac{\omega_n}{\|x_n - x_{n-1}\|}, \alpha_n\} & \text{if } x_n \neq x_{n-1}, \\ \alpha_n & \text{otherwise,} \end{cases} \quad \forall n \in \mathbb{N}$$

or

$$\theta_n = \begin{cases} \begin{cases} \sigma_n \in [0,1) \text{ such that } \sigma_n \to 0 \text{ as } n \to \infty \text{ or} \\ \sigma_n \in [0,1) \text{ such that } \sigma_n \to 1 \text{ as } n \to \infty \text{ or} \\ \sigma_n \in [0,1) \text{ to be chosen arbitrarily} \end{cases} & \text{if } n \le N, \\ \begin{cases} \min\{\frac{\omega_n}{\|x_n - x_{n-1}\|}, \alpha_n\} & \text{if } x_n \neq x_{n-1}, \\ \alpha_n & \text{otherwise,} \end{cases} & \text{otherwise,} \end{cases} \quad \forall n \in \mathbb{N},$$

where $N \in \mathbb{N}$ and $\{\omega_n\}$ is a positive sequence such that $\omega_n = o(\alpha_n)$.

4 IFISTA

Let $F : H \to \mathbb{R}$ be a convex and differentiable function and $G : H \to \mathbb{R} \cup \{+\infty\}$ be a convex and lower semi-continuous function such that the gradient ∇F is L-Lipschitz continuous and ∂G is the subdifferential of G. It is well known that if ∇F is L-Lipschitz continuous, then it is $\frac{1}{L}$-inverse strongly monotone [32]. Moreover, ∂G is maximal monotone [33]. Putting $A = \nabla F$, $B = \partial G$, and $\alpha = \frac{1}{L}$ into Theorem 3.1, we obtain the following result.

Theorem 4.1 *Let H be a real Hilbert space. Let $F : H \to \mathbb{R}$ be a convex and differentiable function with L-Lipschitz continuous gradient ∇F and $G : H \to \mathbb{R}$ be a convex and lower semi-continuous function. Let f be a k-contraction mapping of H into itself, and assume that $(\nabla F + \partial G)^{-1}(0)$ is nonempty. Let $x_0, x_1 \in H$ and $\{x_n\} \subset H$ be a sequence generated by*

$$\begin{cases} z_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = \alpha_n f(z_n) + (1 - \alpha_n) prox_{\lambda_n G}(z_n - \lambda_n \nabla F(z_n) + e_n), \\ x_{n+1} = prox_{\lambda_n G}(y_n - \lambda_n \nabla F(y_n) + e_n), \end{cases}$$

for all $n \in \mathbb{N}$, where $\{\alpha_n\} \subset (0,1)$, $\{\lambda_n\} \subset (0, \frac{2}{L})$, $\{e_n\} \subset H$, and $\{\theta_n\} \subset [0,\theta]$ such that $\theta \in [0,1)$ satisfy the following conditions:

(C1) $\lim_{n\to\infty} \alpha_n = 0$ *and* $\sum_{n=1}^{\infty} \alpha_n = \infty$,

(C2) $0 < a \le \lambda_n \le b < \frac{2}{L}$ *for some* $a, b > 0$,

(C3) $\lim_{n\to\infty} \frac{\|e_n\|}{\alpha_n} = 0$,

(C4) $\sum_{n=1}^{\infty} \|e_n\| < \infty$ *and* $\lim_{n\to\infty} \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| = 0$, *then the sequence* $\{x_n\}$ *converges strongly to a point* $x^* \in (\nabla F + \partial G)^{-1}(0)$, *where* $x^* = P_{(\nabla F+\partial G)^{-1}(0)} f(x^*)$.

We focus on the image restoration using the fixed point optimization algorithm in Theorem 4.1. The image deblurring problem is in the form

$$Ax = b + \varepsilon, \tag{4.1}$$

where $A \in \mathbb{R}^{m\times n}$ represents a known blurring operator (which is called the point spread function: PSF), $b \in \mathbb{R}^m$ is a known observed blurred (and additive noisy) image, $\varepsilon \in \mathbb{R}^m$ is an unknown additive white Gaussian noise, and $x \in \mathbb{R}^n$ is an unknown signal/image to be restored (estimated). Both b and x are formed by stacking the columns of their corresponding two-dimensional images.

In order to solve problem (4.1), we introduce the least absolute shrinkage and selection operator (LASSO) of Tibshirani [34] for solving the following minimization problem:

$$\min_{x\in\mathbb{R}^n} \bigl\{ \|Ax - b\|_2^2 + \lambda \|Wx\|_1 \bigr\}, \tag{4.2}$$

where $\lambda > 0$ is a regularization parameter and $W : \mathbb{R}^n \to \mathbb{R}^n$ represents the orthogonal or tight frame wavelet synthesis, which is a special case of the convex minimization problem (1.2) when $F(x) = \|Ax - b\|_2^2$ and $G(x) = \lambda \|Wx\|_1$ such that $\|x\|_1 = \sum_{i=1}^n |x_i|$ and $\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$ for all $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$. It is well known from Lemma 2.2 by putting $T(Ax) = P_{\mathbb{R}^m} Ax = b$ that $\nabla F(x) = 2A^T(Ax - b)$ and ∇F is L-Lipschitzian with $L = 2\|A\|^2$ such that A^T stands the transpose of A, and $\|A\|$ is the largest singular value of A (i.e., the square root of the largest eigenvalue of the matrix A^TA) or the spectral norm $\|A\|_2$.

In this image deblurring case using Theorem 4.1, if the blurring operator A is smaller than the observed blurred image b and the restored image x, then it is changed by padPSF in MATLAB to embed its array to the matrix $A_{\text{big}} \in \mathbb{R}^{m\times n}$, and followed by a transformation to the signal matrix $A_{\text{sig}} \in \mathbb{R}^{m\times n}$ for calculating the matrix $A_{\text{eig}} = (a_{ij}^{\text{eig}}) \in \mathbb{R}^{m\times n}$ of eigenvalues of the signal matrix A_{sig} using the discrete fast Fourier transformation (FFT) or the discrete cosine transformation (DCT). That is,

$$L = 2\|A_{\text{eig}}\|_{\max}^2 = 2\Bigl(\max_{ij} \bigl|a_{ij}^{\text{eig}}\bigr|\Bigr)^2.$$

We set $m = n$ and the process of gradient ∇F always maps the signal x to 2 times of the signal $A^T(Ax - b)$, where x, A^T, A, and b are in the form of the signal transformation FFT or DCT. That is,

$$\nabla F(x) := \nabla F(x_{\text{sig}}) = 2A_{\text{eig}}^T (A_{\text{eig}} x_{\text{sig}} - b_{\text{sig}}) := 2 \underbrace{A^T(Ax - b)}_{\text{signal form in } \mathbb{R}^m},$$

where $A_{\text{eig}}^T = A_{\text{eig}}^{-1}$ such that A_{eig}^T and A_{eig}^{-1} stand for the transpose and the inverse signal transform of the eigenvalues matrix A_{eig}, respectively.

By [35] and the reference therein, for all $u = (u_1, u_2, \dots, u_m)^T \in \mathbb{R}^m$ and for each $n \in \mathbb{N}$, we have

$$\operatorname{prox}_{\lambda_n G}(u) = \operatorname{prox}_{\lambda_n \lambda \|Wu\|_1}(u) = v$$

Algorithm 1 Improved fast iterative shrinkage thresholding algorithm (IFISTA).

procedure IFISTA

Choose the initials $x_0, x_1 \in \mathbb{R}^m$ arbitrarily.

Set M is the maximum loops to stop and tol is a prescribed tolerance value.

Set the operator A and the mapping f in backing tracks.

$n \leftarrow 0$

repeat

$n \leftarrow n+1$

Update the parameters α_n, λ_n, and θ_n, and the error data $e_n \in \mathbb{R}^m$.

$$z_n \leftarrow x_n + \theta_n(x_n - x_{n-1}),$$

$$y_n \leftarrow \alpha_n f(z_n) + (1-\alpha_n)W^{-1}\big(\mathrm{prox}_{\lambda_n\lambda\|W(\cdot)\|_1}\big(z_n - 2\lambda_n \underbrace{A^T(Az_n - b)}_{\text{signal form in } \mathbb{R}^m} + e_n\big)\big),$$

$$x_{n+1} \leftarrow W^{-1}\big(\mathrm{prox}_{\lambda_n\lambda\|W(\cdot)\|_1}\big(y_n - 2\lambda_n \underbrace{A^T(Ay_n - b)}_{\text{signal form in } \mathbb{R}^m} + e_n\big)\big)$$

until ($n = M$ or $\frac{\|x_{n+1}-x_n\|_2}{\|x_n\|_2} \le \mathrm{tol}$)

return x_{n+1}

end procedure

such that $v = (v_1, v_2, \dots, v_m)^T \in \mathbb{R}^m$, where $v_i = \mathrm{sign}((Wu)_i)\max\{|(Wu)_i| - \lambda_n\lambda, 0\}$ for all $i = 1, 2, \dots, m$. When the process of $\mathrm{prox}_{\lambda_n G}$ has been finished, it returns to $W^{-1}(\mathrm{prox}_{\lambda_n G}(u))$, where W^{-1} stands for the inverse wavelet synthesis such that $W^{-1}(\cdot) = W^T(\cdot)$ before continuing other processes. That is,

$$\mathrm{prox}_{\lambda_n G}\big(z_n - \lambda_n \nabla F(z_n) + e_n\big) = W^{-1}\big(\mathrm{prox}_{\lambda_n\lambda\|W(\cdot)\|_1}\big(z_n - 2\lambda_n \underbrace{A^T(Az_n - b)}_{\text{signal form in } \mathbb{R}^m} + e_n\big)\big)$$

and

$$\mathrm{prox}_{\lambda_n G}\big(y_n - \lambda_n \nabla F(y_n) + e_n\big) = W^{-1}\big(\mathrm{prox}_{\lambda_n\lambda\|W(\cdot)\|_1}\big(y_n - 2\lambda_n \underbrace{A^T(Ay_n - b)}_{\text{signal form in } \mathbb{R}^m} + e_n\big)\big)$$

for all $n \in \mathbb{N}$.

In the next section, we present IFISTA in Algorithm 1 to the improved fast iterative shrinkage thresholding algorithm [19] in the same programming techniques [36].

5 Applications and numerical examples

In this section, we illustrate the performance of IFISTA compared with IFBS, FISTA, FBMWA, FBMMMA, NAGA, and SCTIP for solving the image deblurring problem (4.1) through LASSO problem (4.2) with $\lambda = 10^{-4}$. We implemented them in MATLAB R2019a to solve and run on personal laptop Intel(R) Core(TM) i5-8250U CPU @1.80 GHz 8 GB RAM. We use the quality measures (it is better if it is larger value) of the image restoration as follows.

Let $x, x_n \in \mathbb{R}^{M\times N}$ represent the original image and the estimated image at n^{th} iteration(s), respectively.

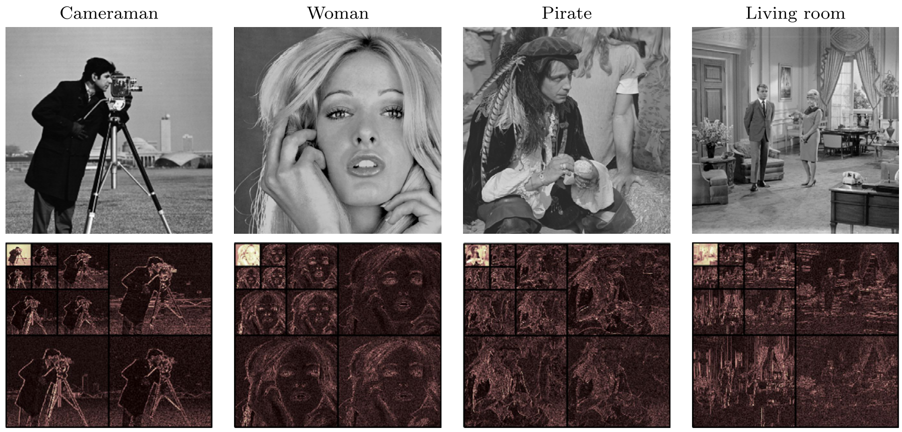

Figure 1 Original images and their 2D three-stage Haar wavelet transform

(1) For looking at how strong the signal is and how strong the noise is, the measure is the signal-to-noise ratio (SNR) of the images x and x_n, which is defined (measured in decibels: dB) by

$$\mathrm{SNR}(x, x_n) = 10 \log_{10} \frac{\|x_n\|_2^2}{\|x - x_n\|_2^2}.$$

(2) For looking at how signal peak is, the measure is the peak signal-to-noise ratio (PSNR) of the images x and x_n, which is defined (measured in decibels: dB) by

$$\mathrm{PSNR}(x, x_n) = 10 \log_{10} \frac{\mathrm{MAX}^2}{\mathrm{MSE}(x, x_n)} = 10 \log_{10} \frac{\mathrm{MAX}^2}{\frac{1}{cMN}\|x - x_n\|_2^2}$$

where MAX is the maximum possible pixel value of the m-unit class (m-bit) image such that $\mathrm{MAX} = 2^m - 1$ (for instance, $\mathrm{MAX} = 255$ for 8-bits image and $\mathrm{MAX} = 65535$ for 16-bits image), and $\mathrm{MSE}(x, x_n)$ is the mean squared error (MSE) of the images x and x_n, which is defined by $\mathrm{MSE}(x, x_n) = \frac{1}{cMN}\|x - x_n\|_2^2$ such that the images x and x_n are c-multichannel image (for instance, $c = 1$ for gray or monochrome image, $c = 3$ for RGB color image, and $c = 4$ for CMYK color image).

Similarly, this measure is the improvement in signal-to-noise ratio (ISNR) of the images x, x_n, and b where the image $b \in \mathbb{R}^{M \times N}$ represents the observed blurred (and additive noisy) c-multichannel image, which is defined (measured in decibels: dB) by

$$\begin{aligned}
\mathrm{ISNR}(x, x_n, b) &= \mathrm{PSNR}(x, x_n) - \mathrm{PSNR}(x, b) \\
&= 10 \log_{10} \frac{\mathrm{MAX}^2}{\frac{1}{cMN}\|x - x_n\|_2^2} - 10 \log_{10} \frac{\mathrm{MAX}^2}{\frac{1}{cMN}\|x - b\|_2^2} \\
&= 10 \log_{10} \frac{\|x - b\|_2^2}{\|x - x_n\|_2^2}.
\end{aligned}$$

For comparison, we consider the standard test images downloaded from [37] for Cameraman, Woman, Pirate, and Living room, with each monochrome images consisting of 512×512 pixels, which represent the original images $x \in \mathbb{R}^{512 \times 512}$, and we converted them

Figure 2 Observed blurred and noisy images

Table 1 The best choice types of testing the parameter λ_n for the fast convergence

Type	A1	A2	A3	A4	A5	B1	B2	C1	C2
λ_n	$\frac{M}{10}$	$M-\frac{M}{10}$	M	$M+\frac{M}{10}$	$2M-\frac{M}{10}$	$\frac{Mn}{n+1}$	$\frac{M(n+2)}{n+1}$	$M+\frac{(-1)^n M}{n+1}$	$M+\frac{(-1)^{n+1} M}{n+1}$

to the double class type by im2double(imread('image_name')) in MATLAB, and also its 2D three-stage Haar wavelet transform $Wx \in \mathbb{R}^{512\times 512}$ as in Fig. 1.

The original images went through a Gaussian blur of size 9×9 and standard deviation 4 as a point spread function (PSF) which represents the blurring operator A by fspecial or psfGauss in MATLAB, and went through imfilter in MATLAB (computed by mirror-reflecting as the array across-the array border or symmetric) and followed by an additive zero-mean white Gaussian noise with standard deviation 10^{-3}, which represents the observed blurred and noisy image $b \in \mathbb{R}^{512\times 512}$ as in Fig. 2. The PSF A was changed by padPSF in MATLAB to embed its array to the matrix $A_{\mathrm{big}} \in \mathbb{R}^{512\times 512}$, and it transformed to a signal matrix $A_{\mathrm{sig}} \in \mathbb{R}^{512\times 512}$ for calculating the matrix $A_{\mathrm{eig}} = (a_{ij}^{\mathrm{eig}}) \in \mathbb{R}^{512\times 512}$ of eigenvalues of the signal matrix A_{sig} using the discrete cosine transformation (DCT). That is, $L = 2\|A_{\mathrm{eig}}\|_{\max}^2 = 2(\max_{ij} |a_{ij}^{\mathrm{eig}}|)^2$.

In compared algorithms, all parameters have been set to their high performance. For each $n \in \mathbb{N}$, we set

$$\alpha_n = \begin{cases} \frac{10^{-6}}{n+1} & \text{for FBMWA, NAGA, SCTIP, IFISTA,} \\ \frac{1}{2}(1-\frac{10^{-6}}{n+1}) & \text{for FBMMMA,} \end{cases}$$

$$\beta_n = \begin{cases} \frac{2n}{5n+1} & \text{for SCTIP,} \\ \frac{10^{-6}}{n+1} & \text{for FBMWA,} \\ \frac{1}{2}(1-\frac{10^{-6}}{n+1}) & \text{for FBMMMA,} \end{cases} \qquad \gamma_n = \begin{cases} 1-10^{-6}-\frac{10^{-6}}{n+1} & \text{for FBMWA,} \\ 1-\frac{10^{-6}}{n+1} & \text{for FBMMMA,} \end{cases}$$

and by [35], we introduced the best choice types of testing the parameter λ_n for the fast convergence as in Table 1 (see also, Tables 1–4 of Examples 4.3, 4.5, 4.8, and 4.10 in [35], respectively) such that $M = \frac{1}{L}$ of $\frac{1}{M}$-Lipschitzian continuous gradient ∇F, it follows the setting to its high performance that

$$\lambda_n = \begin{cases} 2M-\frac{M}{10} & \text{for IFBS, FBMWA, FBMMMA, SCTIP, IFISTA (A5 type),} \\ \frac{M(n+2)}{n+1} & \text{for NAGA (B2 type),} \end{cases}$$

Table 2 The maximum of SNR and ISNR values in first 1^{st} to 100^{th} iteration(s) for image deblurring

Algorithm	*n*	SNR	ISNR	CPU (s)	*n*	SNR	ISNR	CPU (s)
	Cameraman				Woman			
IFBS	80	27.8726	8.4019	7.62	81	24.9248	4.6721	7.71
FISTA	100	27.8425	8.3726	9.44	100	24.9091	4.6573	9.40
FBMWA	56	27.8789	8.4082	14.56	56	24.9239	4.6713	14.58
FBMMMA	49	27.8802	8.4096	11.55	50	24.9239	4.6712	11.72
NAGA	100	27.8549	8.3850	15.44	100	24.9142	4.6623	15.44
SCTIP	71	25.1638	5.6916	8.62	73	23.4103	3.1676	8.72
IFISTA	56	27.8959	8.4252	8.63	56	24.9251	4.6725	8.72
Processing mean	74	27.4841	8.0134	10.84	74	24.7045	4.4535	10.90
Algorithm	*n*	SNR	ISNR	CPU (s)	*n*	SNR	ISNR	CPU (s)
	Pirate				Living room			
IFBS	77	23.7652	5.0117	7.40	80	24.7236	6.7632	7.63
FISTA	100	23.7587	5.0060	9.50	100	24.7026	6.7432	9.30
FBMWA	54	23.7636	5.0101	14.02	55	24.7209	6.7606	14.14
FBMMMA	47	23.7637	5.0104	11.15	49	24.7206	6.7603	11.65
NAGA	100	23.7620	5.0091	15.34	100	24.7097	6.7502	15.35
SCTIP	71	22.1920	3.4440	8.51	75	22.3615	4.4151	8.94
IFISTA	54	23.7607	5.0073	8.34	56	24.7130	6.7526	8.79
Processing mean	72	23.5380	4.7855	10.61	74	24.3788	6.4207	10.83

$$\theta_n = \begin{cases} \sigma_n = \frac{t_n - 1}{t_{n+1}} \text{ such that } t_1 = 1 \text{ and } t_{n+1} = \frac{1+\sqrt{1+4t_n^2}}{2} & \text{if } n \leq 100, \\ \text{(except FISTA, for all } n \in \mathbb{N}) & \\ \begin{cases} \frac{1}{2^n} \text{ for FBMWA, FBMMMA,} & \\ \begin{cases} \min\{\frac{1/(n+1)^2}{\|x_n - x_{n-1}\|_2^2}, 0.5\} & \text{if } x_n \neq x_{n-1}, \\ 0 & \text{otherwise,} \end{cases} & \text{for IFBS} \\ \begin{cases} \min\{\frac{1/(n+1)^3}{\|x_n - x_{n-1}\|_2}, \alpha_n\} & \text{if } x_n \neq x_{n-1}, \\ \alpha_n & \text{otherwise,} \end{cases} & \text{otherwise,} \end{cases} & \text{otherwise,} \end{cases}$$

and the error sequence $\{e_n\} \subset \mathbb{R}^{512\times 512}$ such that

$$e_n = \begin{cases} \frac{b}{n^n} & \text{if } n \leq 100, \\ \frac{b}{(n+1)^3} & \text{otherwise.} \end{cases}$$

We also set $f(x) = \frac{x}{5}$ for all $x \in \mathbb{R}^{512\times 512}$ and choose the initials $x_0 = x_1 = b$ for all algorithms (except for FISTA, $y_1 = x_0 = b$). Quoting from [38], we can use max-norm regularization, this constrains the norm of the vector of incoming weights at each hidden unit to be bound by a constant c. Max-norm regularization was used for weights in both convolutional and fully connected layers. Since $L = 2\|A_{\text{eig}}\|_{\max}^2$, so we can use SNR and ISNR that both are two quality measures of the image restoration (it is better if it is larger value) to find each hidden estimated image before its convergence in first 1^{st} to 100^{th} iteration(s) to show high performance of each compared algorithm. That is, we find the hidden estimated images x^* and y^* such that

$$\text{SNR}(x, x^*) = \max_{1\leq n\leq 100} \text{SNR}(x, x_n) \text{ and } \text{ISNR}(x, y^*, b) = \max_{1\leq n\leq 100} \text{ISNR}(x, x_n, b),$$

Table 3 The SNR and ISNR values at first 1000th iterations for image deblurring

Algorithm	SNR	ISNR	CPU (s)	tol	SNR	ISNR	CPU (s)	tol
	Cameraman				Woman			
IFBS	27.3690	7.8949	94.08	7.52×10^{-6}	24.7271	4.4717	93.94	6.70×10^{-6}
FISTA	25.1667	5.6856	90.29	1.60×10^{-5}	23.7183	3.4576	90.52	1.52×10^{-5}
FBMWA	26.8948	7.4189	247.48	6.65×10^{-6}	24.5520	4.2953	250.39	5.81×10^{-6}
FBMMMA	26.6595	7.1828	227.37	6.47×10^{-6}	24.4583	4.2011	227.31	5.60×10^{-6}
NAGA	27.8742	8.4032	151.89	7.20×10^{-6}	24.9137	4.6609	151.08	6.49×10^{-6}
SCTIP	26.3036	6.8349	118.22	1.64×10^{-5}	23.9936	3.7433	118.25	1.39×10^{-5}
IFISTA	26.8945	7.4186	151.80	6.66×10^{-6}	24.5520	4.2953	152.43	5.83×10^{-6}
Processing mean	26.7375	7.2633	154.45	9.56×10^{-6}	24.4164	4.1607	154.85	8.50×10^{-6}
Algorithm	SNR	ISNR	CPU (s)	tol	SNR	ISNR	CPU (s)	tol
	Pirate				Living room			
IFBS	23.4923	4.7342	93.61	8.29×10^{-6}	24.3377	6.3731	94.02	8.43×10^{-6}
FISTA	22.4952	3.7298	89.79	1.58×10^{-5}	22.5101	4.5352	90.06	2.01×10^{-5}
FBMWA	23.2622	4.5020	250.55	7.14×10^{-6}	23.9751	6.0082	247.41	7.90×10^{-6}
FBMMMA	23.1511	4.3900	225.98	6.83×10^{-6}	23.7977	5.8298	226.78	7.76×10^{-6}
NAGA	23.7434	4.9891	150.45	8.10×10^{-6}	24.6826	6.7220	151.47	8.19×10^{-6}
SCTIP	22.6546	3.9038	117.91	1.71×10^{-5}	23.3893	5.4341	116.67	1.76×10^{-5}
IFISTA	23.2625	4.5022	151.51	7.17×10^{-6}	23.9766	6.0097	152.14	7.94×10^{-6}
Processing mean	23.1516	4.3930	154.26	1.01×10^{-5}	23.8099	5.8446	154.08	1.11×10^{-5}

Table 4 The effectiveness comparison of image deblurring

Algorithm	CPU	SNR	ISNR	*n*
In the fast processing with high performance in first 1st to 100th iteration(s) for image deblurring.				
IFBS	√	√	√	×
FISTA	√	√	√	×
FBMWA	×	√	√	√
FBMMMA	×	√	√	√
NAGA	×	√	√	×
SCTIP	√	×	×	√
IFISTA	√	√	√	√

Algorithm	CPU	SNR	ISNR	*n*	tol
In the fast convergence with good performance at first 1000th iterations for image deblurring.					
IFBS	√	√	√	√	√
FISTA	√	×	×	√	×
FBMWA	×	√	√	√	√
FBMMMA	×	×	×	√	√
NAGA	√	√	√	√	√
SCTIP	√	×	×	√	×
IFISTA	√	√	√	√	√

√: satisfy and ×: unsatisfy (compared with processing mean) (Evaluation order: CPU, SNR, ISNR, *n*, tol.)

it is better if $x^* = y^*$ (which means that both hidden estimated images are in the process of the same iteration), which is shown in Table 2. Moreover, we also show the relative error which is defined by

$$\frac{\|x_{n+1} - x_n\|_2}{\|x_n\|_2} \leq \text{tol},$$

where tol denotes a prescribed tolerance value of each compared algorithm at first 1000th iterations by the constants SNR and ISNR as in Table 3, and their convergence behavior are shown in Fig. 3 and Fig. 4.

In evaluation of each algorithm, we use the image processing mean to assess them such that $\overline{n}$, $\overline{\text{SNR}}$, $\overline{\text{ISNR}}$, $\overline{\text{CPU}}$, and $\overline{\text{tol}}$ are the compared image processing arithmetic mean of n, SNR, ISNR, CPU times, and tol, respectively.

On the results of each algorithms in first 1st to 100th iteration(s) as in Table 2, we see that IFBS, FISTA, and IFISTA have the quantities of SNR and ISNR near to others (except for SCTIP) and their quantities of n and CPU times are vastly different from others. We

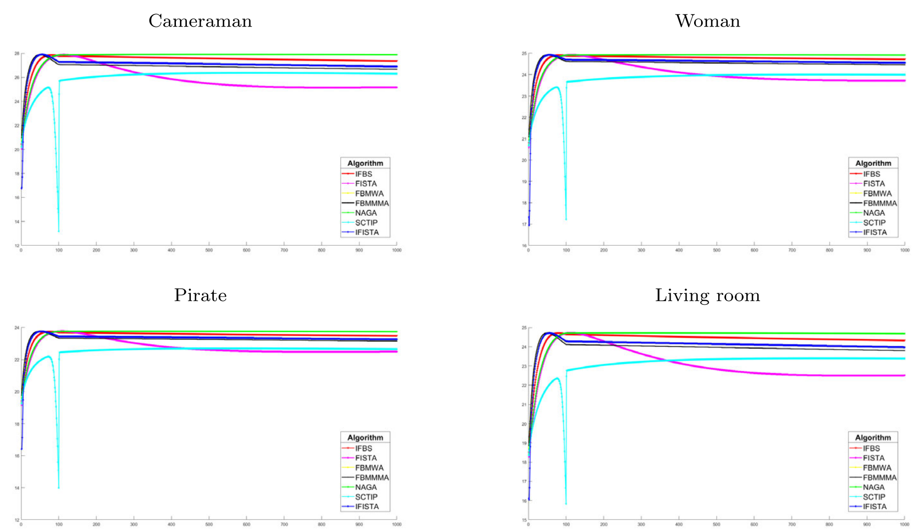

Figure 3 The SNR convergence behavior of image deblurring

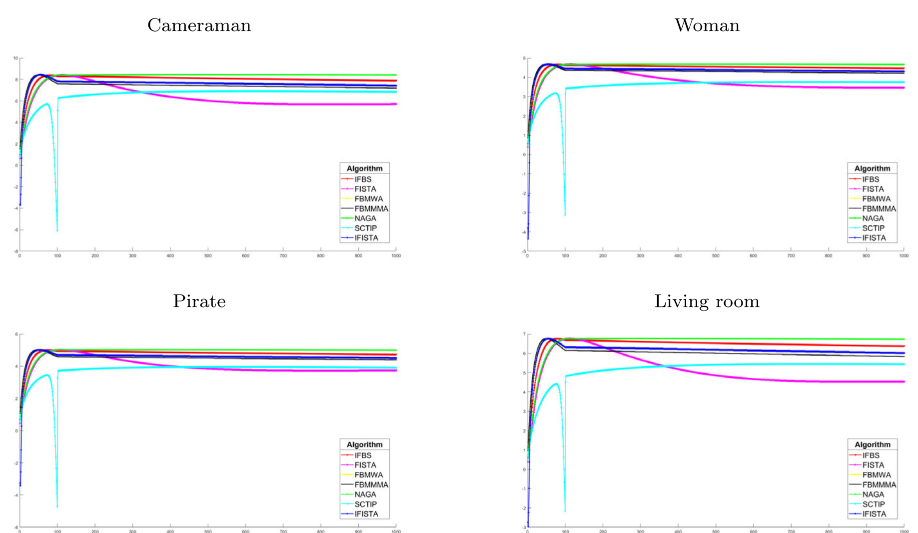

Figure 4 The ISNR convergence behavior of image deblurring

give an evaluation order for those algorithms as in Table 4 as follows: CPU $\leq \overline{\mathrm{CPU}}$, SNR $\geq \overline{\mathrm{SNR}}$, ISNR $\geq \overline{\mathrm{ISNR}}$, and $n \leq \overline{n}$, respectively. We see that all evaluations of IFISTA are satisfied, while only CPU times, SNR, and ISNR evaluations of both IFBS and FISTA are satisfied, where SNR and ISNR of IFBS are greater than FISTA, and then, we conclude that IFISTA, IFBS, and FISTA are in the 1$^{\text{st}}$, 2$^{\text{nd}}$, and 3$^{\text{rd}}$ place, respectively, of the top three fast processing with high performance for compared image deblurring as Fig. 5, Fig. 6, Fig. 7, and Fig. 8.

From results of each algorithms at first 1000$^{\text{th}}$ iterations as in Table 3, we see that the quantities of SNR and ISNR of all algorithms are vastly different. We give an evaluation order for those algorithms as in Table 4 as follows: CPU $\leq \overline{\mathrm{CPU}}$, SNR $\geq \overline{\mathrm{SNR}}$, ISNR $\geq \overline{\mathrm{ISNR}}$, $n \leq \overline{n}$, and tol $\leq \overline{\mathrm{tol}}$, respectively. We see that all evaluations of IFBS, NAGA, and IFISTA

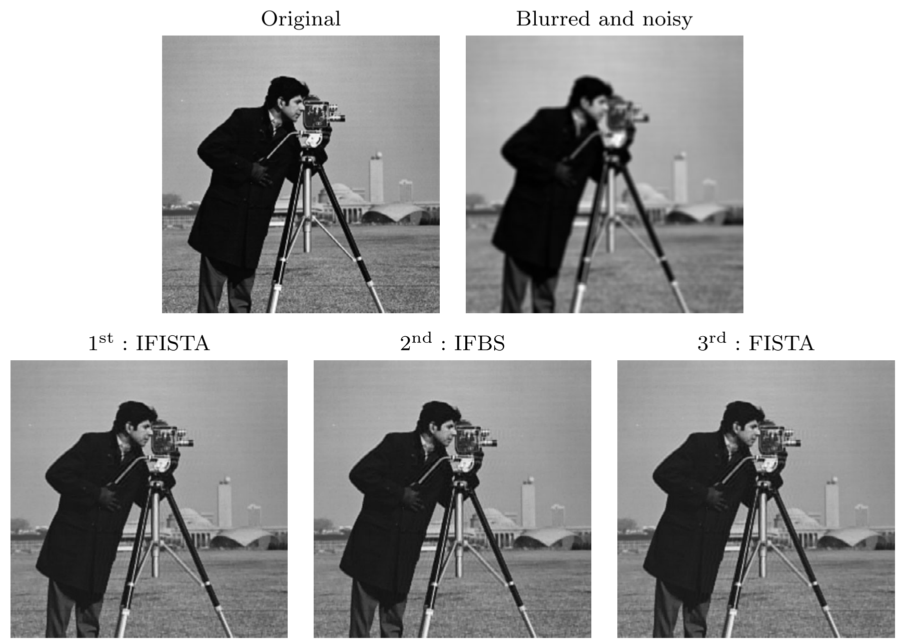

Figure 5 Top three fast processing with high performance for deblurring of Cameraman

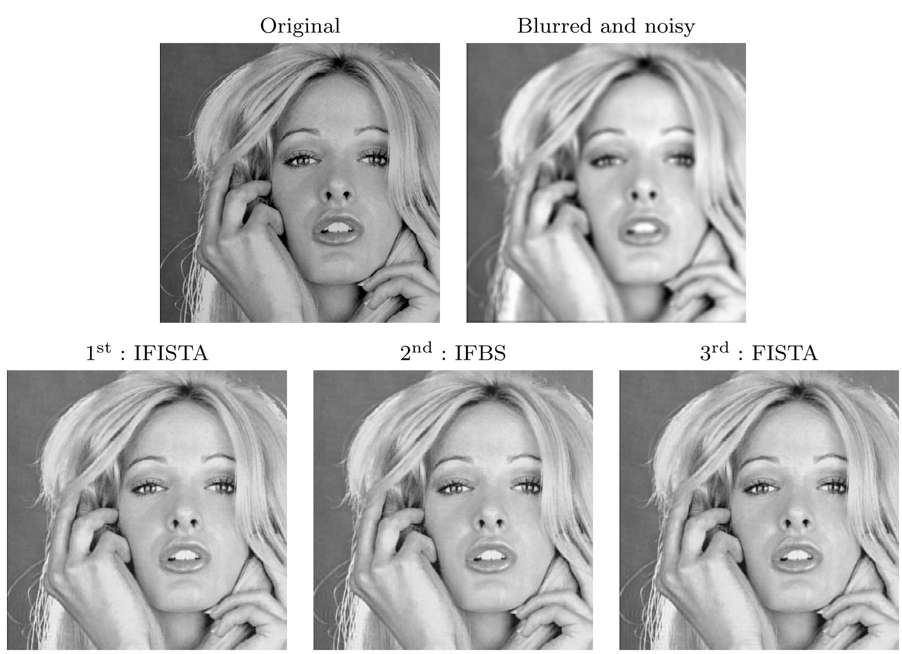

Figure 6 Top three fast processing with high performance for deblurring of Woman

are satisfied, where SNR and ISNR of NAGA are greater than both IFBS and IFISTA, and also which of IFBS are greater than IFISTA, and then, we conclude that NAGA, IFBS, and IFISTA are in the 1^{st}, 2^{nd}, and 3^{rd} place, respectively, of the top three fast convergence with good performance for compared image deblurring as Fig. 9, Fig. 10, Fig. 11, and Fig. 12.

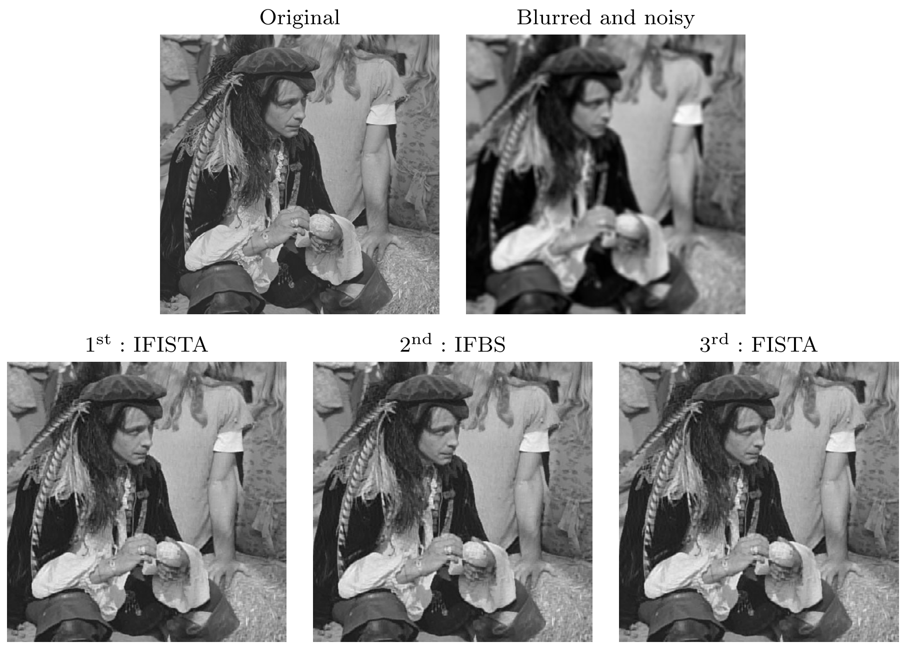

Figure 7 Top three fast processing with high performance for deblurring of Pirate

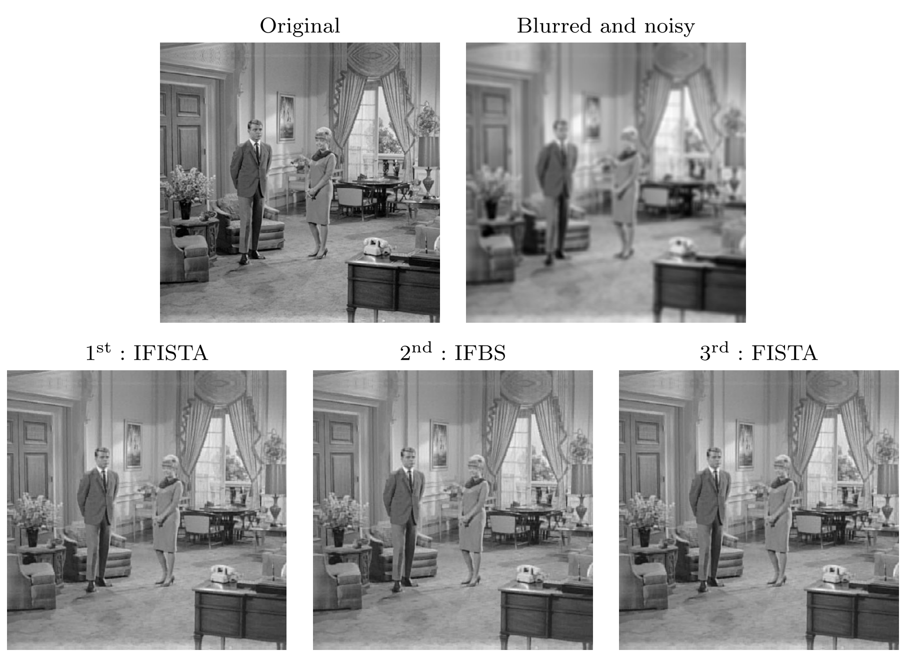

Figure 8 Top three fast processing with high performance for deblurring of Living room

6 Conclusion

A new iterative forward-backward splitting method with an error is obtained in our main result. It can be applied to improve the fast iterative shrinkage thresholding algorithm (IFISTA) with an error for solving the image deblurring problem. Under the same pro-

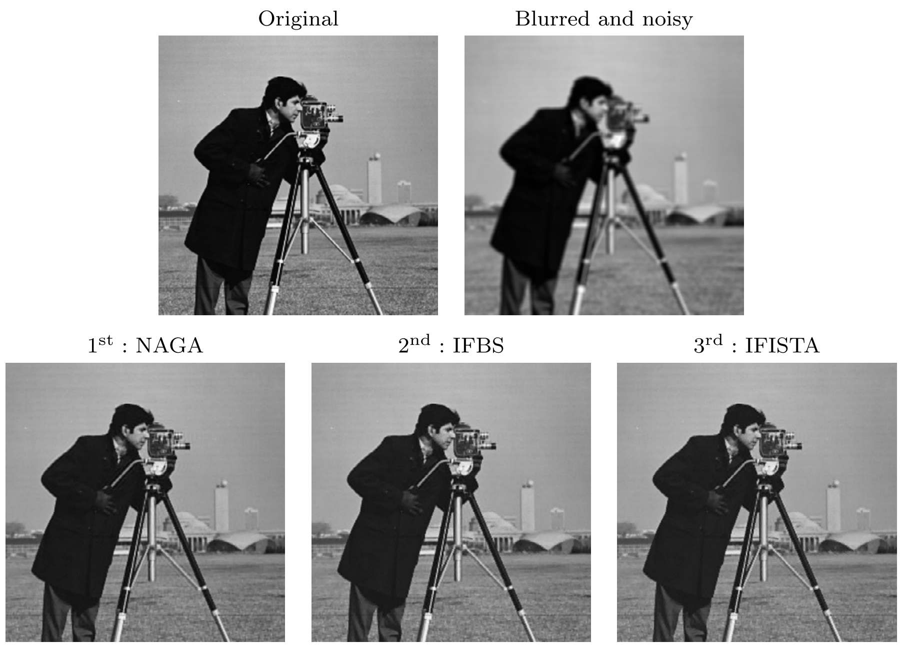

Figure 9 Top three fast convergence with good performance for deblurring of Cameraman

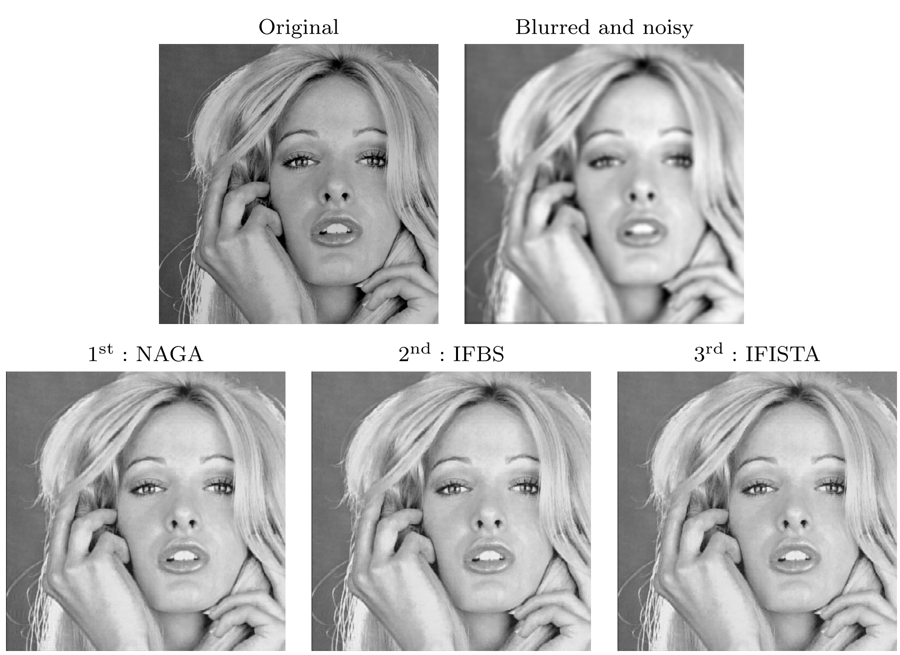

Figure 10 Top three fast convergence with good performance for deblurring of Woman

gramming techniques [36] and setting all parameters to their high performance, we obtain the following results.

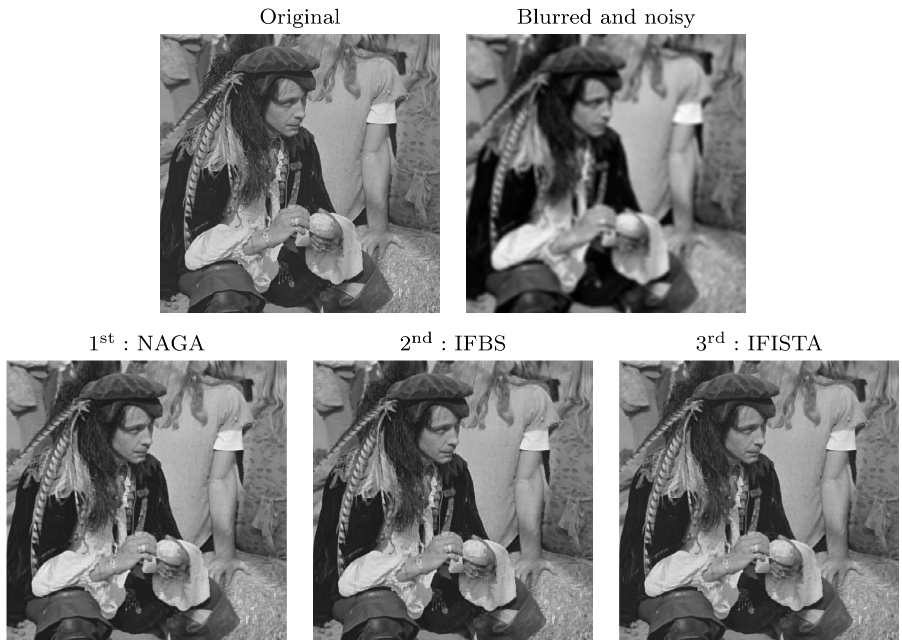

Figure 11 Top three fast convergence with good performance for deblurring of Pirate

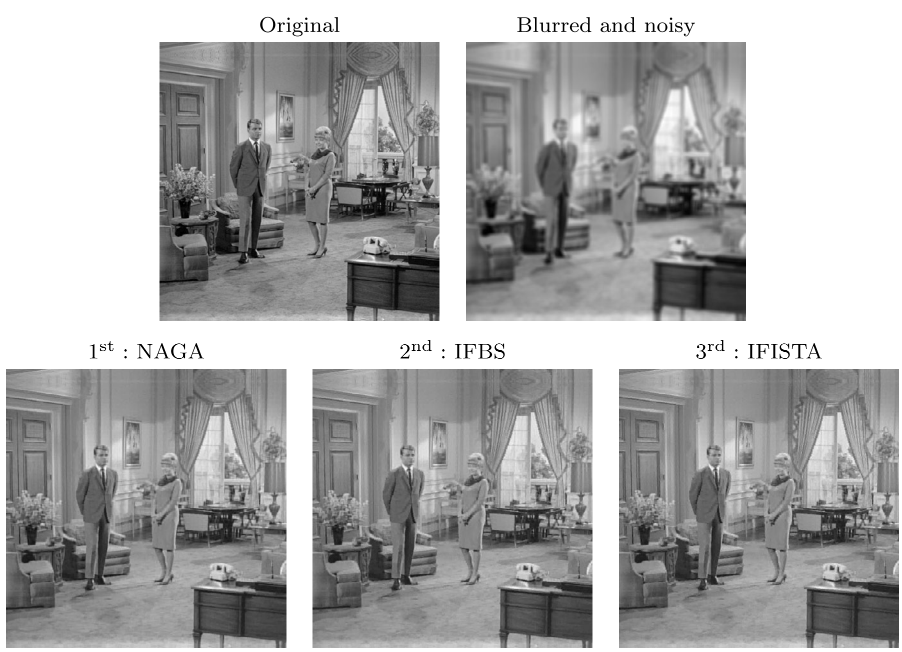

Figure 12 Top three fast convergence with good performance for deblurring of Living room

1. For looking at the fast processing with high performance for compared image deblurring, IFISTA, IFBS, and FISTA are in the 1st, 2nd, and 3rd place, respectively, and all are better than FBMWA, FBMMMA, NAGA, and SCTIP.

2. For looking at the fast convergence with good performance for compared image deblurring, NAGA, IFBS, and IFISTA are in the 1^{st}, 2^{nd}, and 3^{rd} place, respectively, and all are better than FISTA, FBMWA, FBMMMA, and SCTIP.

Acknowledgements
The author would like to thank the Faculty of Science, Maejo University for its financial support.

Funding
This research was supported by Faculty of Science, Maejo University.

Availability of data and materials
Not applicable.

Declarations

Competing interests
The author declares that he has no competing interests.

Authors' contributions
All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Publisher's Note

Received: 4 August 2021 Accepted: 28 September 2021 Published online: 01 November 2021

References
1. Bauschke, H.H.: The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space. J. Math. Anal. Appl. **202**, 150–159 (1996)
2. Chidume, C.E., Bashir, A.: Convergence of path and iterative method for families of nonexpansive mappings. Appl. Anal. **67**, 117–129 (2008)
3. Halpern, B.: Fixed points of nonexpansive maps. Bull. Am. Math. Soc. **73**, 957–961 (1967)
4. Ishikawa, S.: Fixed points by a new iteration method. Proc. Am. Math. Soc. **44**, 147–150 (1974)
5. Klen, R., Manojlovic, V., Simic, S., Vuorinen, M.: Bernoulli inequality and hypergeometric functions. Proc. Am. Math. Soc. **142**, 559–573 (2014)
6. Kunze, H., La Torre, D., Mendivil, F., Vrscay, E.R.: Generalized fractal *n* transforms and self-similar objects in cone metric spaces. Comput. Math. Appl. **64**, 1761–1769 (2012)
7. Mann, W.R.: Mean value methods in iteration. Proc. Am. Math. Soc. **4**, 506–510 (1953)
8. Radenovic, S., Rhoades, B.E.: Fixed point theorem for two non-self mappings in cone metric spaces. Comput. Math. Appl. **57**, 1701–1707 (2009)
9. Todorcevic, V.: Harmonic Quasiconformal Mappings and Hyperbolic Type Metrics. Springer, Basel (2019)
10. Byrne, C.: Iterative oblique projection onto convex subsets and the split feasibility problem. Inverse Probl. **18**, 441–453 (2002)
11. Byrne, C.: A unified treatment of some iterative algorithms in signal processing and image reconstruction. Inverse Probl. **20**, 103–120 (2004)
12. Combettes, P.L., Wajs, V.: Signal recovery by proximal forward-backward splitting. Multiscale Model. Simul. **4**, 1168–1200 (2005)
13. Censor, Y., Bortfeld, T., Martin, B., Trofimov, A.: A unified approach for inversion problems in intensity-modulated radiation therapy. Phys. Med. Biol. **51**, 2353–2365 (2006)
14. Censor, Y., Elfving, T., Kopf, N., Bortfeld, T.: The multiple set split feasibility problem and its applications. Inverse Probl. **21**, 2071–2084 (2005)
15. Censor, Y., Motova, A., Segal, A.: Perturbed projections and subgradient projections for the multiple-sets feasibility problem. J. Math. Anal. **327**, 1244–1256 (2007)
16. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**, 964–979 (1979)
17. Passty, G.B.: Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. J. Math. Anal. **72**, 383–390 (1979)
18. Moudafi, A., Oliny, M.: Convergence of a splitting inertial proximal method for monotone operators. J. Comput. Appl. Math. **155**, 447–454 (2003)
19. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**(1), 183–202 (2009)
20. Hanjing, A., Suantai, S.: A fast image restoration algorithm based on a fixed point and optimization method. Mathematics **8**, 378 (2020)
21. Padcharoen, A., Kumam, P.: Fixed point optimization method for image restoration. Thai J. Math. **18**(3), 1581–1596 (2020)
22. Verma, M., Shukla, K.K.: A new accelerated proximal gradient technique for regularized multitask learning framework. Pattern Recognit. Lett. **95**, 98–103 (2017)

23. Cholamjiak, P., Kesornprom, S., Pholasa, N.: Weak and strong convergence theorems for the inclusion problem and the fixed-point problem of nonexpansive mappings. Mathematics **7**, 167 (2019)
24. Abubakar, J., Kumam, P., Ibrahim, A.H., Padcharoen, A.: Relaxed inertial Tseng's type method for solving the inclusion problem with application to image restoration. Mathematics **8**, 818 (2020)
25. Luo, Y.: An inertial splitting algorithm for solving inclusion problems and its applications to compressed sensing. J. Appl. Numer. Optim. **2**(3), 279–295 (2020)
26. Takahashi, W.: Introduction to Nonlinear and Convex Analysis. Yokohama Publishers, Yokohama (2009)
27. Tang, J.F., Chang, S.S., Yuan, F.: A strong convergence theorem for equilibrium problems and split feasibility problems in Hilbert spaces. Fixed Point Theory Appl. **2014**, 36 (2014)
28. Nadezhkina, N., Takahashi, W.: Weak convergence theorem by an extragradient method for nonexpansive mappings and monotone mappings. J. Optim. Theory Appl. **128**, 191–201 (2006)
29. Geobel, K., Kirk, W.A.: Topic in Metric Fixed Point Theory. Cambridge Studies in Advanced Mathematics, vol. 28. Cambridge University Press, Cambridge (1990)
30. Takahashi, W., Xu, H.-K.: Iterative algorithms for nonlinear operators. J. Lond. Math. Soc. **66**, 240–256 (2002)
31. He, S., Yang, C.: Solving the variational inequality problem defined on intersection of finite level sets. Abstr. Appl. Anal. **2013**, Article ID 942315 (2013)
32. Baillon, J.B., Haddad, G.: Quelques proprietes des operateurs angle-bornes et cycliquement monotones. Isr. J. Math. **26**, 137–150 (1977)
33. Rockafellar, R.T.: On the maximal monotonicity of subdifferential mappings. Pac. J. Math. **33**, 209–216 (1970)
34. Tibshirani, R.: Regression shrinkage and selection via the lasso. J. R. Stat. Soc., Ser. B, Stat. Methodol. **58**, 267–288 (1996)
35. Tianchai, P.: The zeros of monotone operators for the variational inclusion problem in Hilbert spaces. J. Inequal. Appl. **2021**, 126 (2021)
36. Guide to the MATLAB code for wavelet-based deblurring with FISTA, Available online: https://docplayer.net/128436542-Guide-to-the-matlab-code-for-wavelet-based-deblurring-with-fista.html (accessed on 1 June 2021)
37. Image Databases, Available online: http://www.imageprocessingplace.com/downloads_V3/root_downloads/image_databases/standard_test_images.zip (accessed on 1 June 2021)
38. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**, 1929–1958 (2014)