IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 21, 2019, accepted November 7, 2019, date of publication November 15, 2019,
date of current version November 27, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2953795

Bias Correction of Multiple MRI Images Based on an Improved Nonparametric Maximum Likelihood Method

YAN XU[1,2], SHUNBO HU[2], AND YUYUE DU[1]
[1]College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China
[2]College of Information Science and Engineering, Linyi University, Linyi 276005, China

Corresponding authors: Yuyue Du (yydu001@163.com) and Shunbo Hu (hushunbo@lyu.edu.cn).

This work was supported in part by the National Natural Science Foundation of China under Grant 61771230, Grant 61903229, and Grant 61973180, in part by the Key Research and Development Program of Shandong Province under Grant 2018GGX101011, and in part by the Natural Science Foundation of Shandong Province of China under Grant ZR2016FM40, Grant ZR2018MF001, Grant ZR2019BF004, and Grant ZR2019PF005.

ABSTRACT With the wide application of nuclear magnetic resonance imaging, the multiplicative bias field in nuclear magnetic resonance images has created great difficulties for doctors in reading diagnostics and for computers in autoprocessing. Most previous methods eliminate the bias field in the image by estimating a single unknown bias field. An improved method that uses the nonparametric maximum likelihood to jointly eliminate bias from multiple magnetic resonance imaging (MRI) images is proposed in this paper. The method uses the statistics from the same location across different patient images, rather than within an image, and builds a ''multiresolution'' nonparametric tissue model conditioned on image location. We use a separate and nonparametric model to consider the intensity values at each pixel and utilize nonparametric maximum likelihood distance measures to simultaneously eliminate the bias of magnetic resonance (MR) images from different patients. Finally, the performance of the same was tested on a synthetic MRI dataset and a real MRI dataset and is found that the proposed algorithm provides better performance than the method of using entropy minimization across images and the most popular and widely used method, N4.

INDEX TERMS Magnetic resonance imaging (MRI), bias field, intensity inhomogeneity correction, nonparametric maximum likelihood, distance measure.

I. INTRODUCTION

With the rapid development of medical imaging technology as a non-invasive, non-ionizing radiation diagnostic tool, magnetic resonance imaging (MRI) has increasingly become a commonly used detection method in clinical practice. However, the problem of grayscale in-homogeneity often occurs with MRI. The mean and variance of different local regions of the same tissue have a large bias, which is called the ''in-homogeneous field'' or ''bias field''. Additionally, this problem has become more prominent after the appearance of high field intensity nuclear magnetic resonance equipment. The variation of MRI density caused by the bias field can occasionally reach 30% of the normal image density gain [1], which causes great trouble for doctors in image assessment and clinical diagnosis, and has an adverse effect on computer digital processing of medical images, such as segmentation, registration, and quantification.

The MRI bias field is generally defined as a slowly varying and smooth multiplicative field added to the real undisturbed MRI image. The main factors that cause this multiplicative field are [1], [2] (1) an in-homogeneous scanning magnetic field, (2) a difference in the magnetization characteristics (magnetic susceptibility) of the biological tissue, (3) differential attenuation of signal determined by structure position, (4) radio frequency penetration, (5) induced current due to gradient changes [3] and (6) other nonlinear characteristics of the scanning device (sensitivity of the receiving coil, operating conditions, age and state of the device, etc.) [4], [5]. Under normal circumstances, the intensity non-uniformity is mainly caused by (1), and a low-frequency brightness change occurs between the entire field of view and the slice [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Abdel-Hamid Soliman.

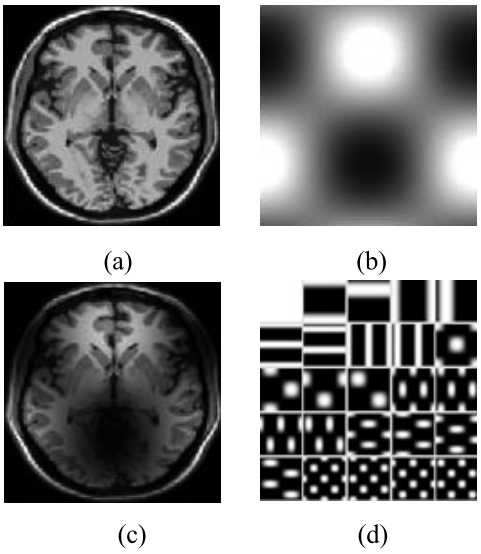

FIGURE 1. A MR image of the human brain contaminated by a multiplicative bias field. (a) An intrinsic image. (b) Simulated low-frequency bias fields. (c) The contaminated image. (d) Two-dimensional sine/cosine bases for constructing band-limited bias fields.

The goal of magnetic resonance (MR) bias correction is to estimate an undamaged image from the corrupted image, as shown in Fig. 1. Preprocessing correction methods based on the characteristics of magnetic resonance imaging equipment have substantial limitations and are not suitable for clinical applications. A large number of image post-processing correction methods have been introduced, and are generally divided into the following [7], [8]: (1) filter-based methods [9], e.g., homomorphic filtering and wavelet-based multiscale filtering algorithms, (2) statistical model-based methods [10], (3) spatial shape information model-based methods, (4) methods based on a surface fitting field [1], [11], and (5) histogram based methods [10], [12]. Wells *et al.* [13] developed a statistical model with discrete organizational sets that uses the expectation maximization (EM) to simultaneously estimate the bias fields, organizational types and residual noise. This approach works well in many cases, but the model must be developed from a ''no bias'' image and the number of tissues is fixed. Viola [14] proposed a nonparametric method, which assumes that he bias correction field would be estimated by minimizing the distribution entropy of pixel brightness. Although this method solves several problems of the fixed tissue parameter model, the statistical model only comes from the data of a single image, which is weak, and no mechanism exists to distinguish selected low-frequency image components from the bias field. This method may mistake the signal for noise in certain cases.

The N4 bias correction method proposed by Tustison *et al.* [12] is currently the most popular and widely used bias correction method, which used robust B-spline approximation routine and a modified hierarchical optimization scheme for bias field correction [15]. N4 method requires foreground extraction and efficiency might depend on the accuracy of background removal, initialization of bias field histogram width and spline distance [16]. It is time consuming when N4 is implemented in The Insight Segmentation and Registration Toolkit (ITK).

Erik proposed an algorithm for joint removal of MRI bias using entropy minimization across images [17]–[19]. The algorithm utilizes data from multiple MR images of different patients to provide improved distribution estimation, establishes a ''multiresolution'' nonparametric tissue model based on image position, and uses entropy minimization measures to remove bias across images. This approach reduces the need for unbiased images in model creation. However, in the organization model, the image entropy and entropy gradient are calculated by using the log function multiple times across the images, and as the number of images increases, the bias removal effect improves, but the calculation also increases sharply. If the number of images is small and the bias is serious, the model removes the bias effect poorly. Therefore, we change the intensity similarity from cross-image minimum entropy to cross-image minimum distance in this paper. The improved method can greatly reduce the calculation amount and enhance the calculation speed in the case of a large number of images, and can achieve a better bias removal effect when the number of images is small and the image corruption is serious.

II. PROBLEM DESCRIPTION AND IMAGE MODEL

It is assumed that each image $I^i\,(1 \le i \le N)$ in a brain MR image set $\boldsymbol{I}\left(I^1, I^2, \ldots, I^N\right)$ contaminated by a bias field (called observation image set) obtained from a fixed population can be approximated by a multiplicative field [6], which can be expressed by the following mathematical model:

$$I^i(x,y) = L^i(x,y)^* B^i(x,y) + NO(x,y), \tag{1}$$

where $L^i\,(1 \le i \le N) \in L$ is a bias-free image (also called intrinsic image); $B^i\,(1 \le i \le N) \in B$ is a smooth bias field; $NO\,(x, y)$ is additive white Gaussian noise or an additive bias field, which can be removed by conventional low frequency filtering; $\boldsymbol{L}$ is the intrinsic image set; $\boldsymbol{B}$ is the set of unknown smooth bias fields; (x, y) is the pixel position, and $\boldsymbol{N}$ is the total number of images. Therefore, formula (1) can be simplified as follows:

$$I^i(x,y) = L^i(x,y)^* B^i(x,y). \tag{2}$$

Fig. 1(a) is a bias-free and noiseless intrinsic image L^i obtained from the Brain Web Simulated Brain Database [20], Fig. 1(b) is a low-frequency bias field B^i constructed by Fig. 1(d), and Fig. 1(c) is the contaminated image I^i obtained by multiplying Fig. 1(a) and Fig. 1(b). The goal of MR bias correction is to recover Fig. 1(a) from Fig. 1(c).

Pixels of specific locations of each image in the contaminated image set (e.g., portions indicated by the small squares in Fig. 2) are aggregated to form a pixel stack. Fig. 2(c) is the histogram distribution of the pixel stack formed by the small squares in Fig. 2(a), and Fig. 2(d) is the histogram distribution of the pixel stack formed by the small squares

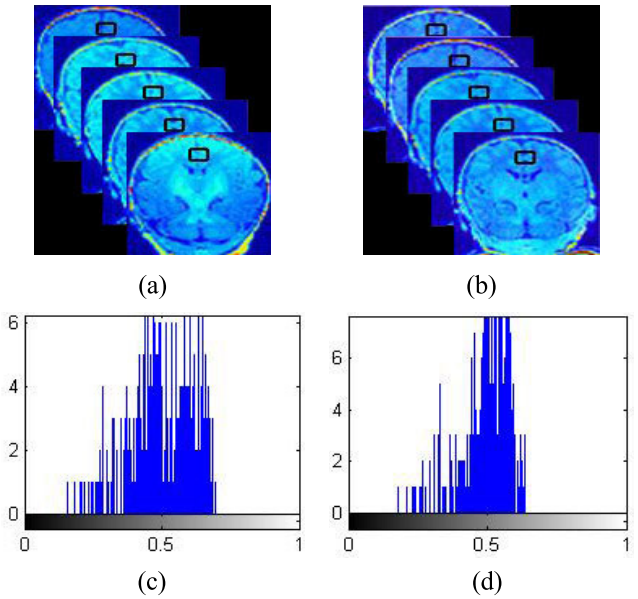

FIGURE 2. The distribution changes of gray histogram. (a) Contaminated brain images from five different infants. (b) Corrected images using our bias field removal algorithm. (c) Histogram distribution of pixels in the pixel stack composed of small squares in (a). (d) Histogram distribution of pixels in the pixel stack composed of the same small squares in (b).

in Fig. 2(b). According to the distribution of the histogram, due to the nonlinear mapping of the bias field in Fig. 2(a), the intensity of the tissue originally distributed in a narrow band is widened, and the intensity distribution of each tissue is severely overlapped, which makes the boundaries between the peaks non-obvious, and the distribution of intensity tends to be uniform. This nonlinear mapping leads to an extension of the grayscale distribution of each type of tissue, which causes an increase in the image entropy.

The method of Erik [17] is based on the principle that when the bias fields have been removed, the pixel-stack values are similar, and in other words, they have lower empirical entropy. This method assumes that all pixels in the intrinsic image generation model are independent, the probability distribution of the intensity of each pixel at the location (x, y) is $p_{x,y}(\cdot)$, and the bias fields of each image are chosen independently and randomly. Unlike most models that rely on statistical rules within an image, this model uses a completely orthogonal approach by assuming that pixel values are independent given their image locations, and when the bias field is removed, the pixel stacks in general have low entropy.

Erik [17] expressed the problem as a maximum a posteriori (MAP) problem that searches for the most likely bias field given the set of contaminated images $\boldsymbol{I}$. Let $\mathcal{B}$ denote the space composed of a two-dimensional sine/cosine base of the smoothed bias fields (corresponding to the 25 base images of Fig. 1(d)).

$$\begin{aligned}&\arg\max_{B\in\mathcal{B}} P\left(\mathrm{B}\,|\mathrm{I}\right)\\ &\overset{(a)}{\approx} \arg\min_{B\in\mathcal{B}} \sum_{x,y} H\left(p_{x,y}\right)\\ &\overset{(b)}{\approx} \arg\min_{B\in\mathcal{B}} \sum_{x,y} \hat{H}_{Vasicek}\left(L^1(x,y),\ldots,L^N(x,y)\right)\\ &= \arg\min_{B\in\mathcal{B}} \sum_{x,y} \hat{H}_{Vasicek}\left(\frac{I^1(x,y)}{B^1(x,y)},\ldots,\frac{I^N(x,y)}{B^N(x,y)}\right). \quad (3)\end{aligned}$$

In Formula (3), H is the Shannon entropy [21], and $\hat{H}_{Vasicek}$ is a sample-based entropy estimator (See Formula (4)). Vasicek's [22] estimator can quickly calculate the entropy from many samples, instead of estimating the distribution from the data estimate entropy from samples or instead of estimating the distribution from the data. In Formula (3), (a) is the approximate value of image entropy, and (b) is the estimation of the image entropy directly from the samples in the pixel stack by the entropy estimator $\hat{H}_{Vasicek}$ without calculating $p_{x,y}(\cdot)$. When N is large enough, i.e., when the number of images is large enough, the inequalities (a) and (b) become equality in Formula (3), and $\hat{H}_{Vasicek}$ is calculated as follows:

$$\begin{aligned}&\hat{H}_{\mathrm{vasicek}}\left(Z^1,\ldots,Z^N\right)\\ &\quad= \frac{1}{N-m}\sum_{i=1}^{N-m}\log\left(\frac{N}{m}\left(Z^{(i+m)}-Z^{(i)}\right)\right), \quad (4)\end{aligned}$$

where $Z^{i'}s$ represent the values in a pixel-stack, $Z^{(i)}$ represents those same values in rank order, N represents the number of values in the pixel-stack, and m is a function of N (usually taken as $m = \sqrt{N}$). For example, when $m = 1$, $Z^{(i+m)} - Z^{(i)}$ is the distance between two adjacent levels, and the estimated result is proportional only to the sum of the logarithm of these distances. These entropy estimators are discussed at length elsewhere [23].

If the intensities of many pixels are concentrated at the same region, the distance among these intensities is small, that is, the intensity variations in the pixel stack are small. If the distribution of intensity is diverged, the distance among these pixels becomes larger. Therefore, to improve the calculation efficiency and the correction effect of the bias field at the same time, the logarithm of the distance in Formula (4) is directly replaced by the 1-norm or 2-norm function. In this paper, we use the Manhattan distance (1-norm) or the squared Euclidean distance (2-norm), and the distance estimators are as described as follows:

$$\hat{D}_{\mathrm{Manhattan}}\left(Z^1,\ldots,Z^N\right)=\frac{1}{N-m}\sum_{i=1}^{N-m}\left|\frac{N}{m}\left(Z^{(i+m)}-Z^{(i)}\right)\right|. \quad (5)$$

$$\hat{D}_{\mathrm{SEuclidean}k}\left(Z^1,\ldots,Z^N\right)=\frac{1}{N-m}\sum_{i=1}^{N-m}\frac{N}{m}\left(Z^{(i+m)}-Z^{(i)}\right)^2. \quad (6)$$

III. ALGORITHM DESCRIPTIONS

The linear model of the bias fields is chosen to optimize Equation (5) or Equation (6) over the set of a 2-D sine/cosine

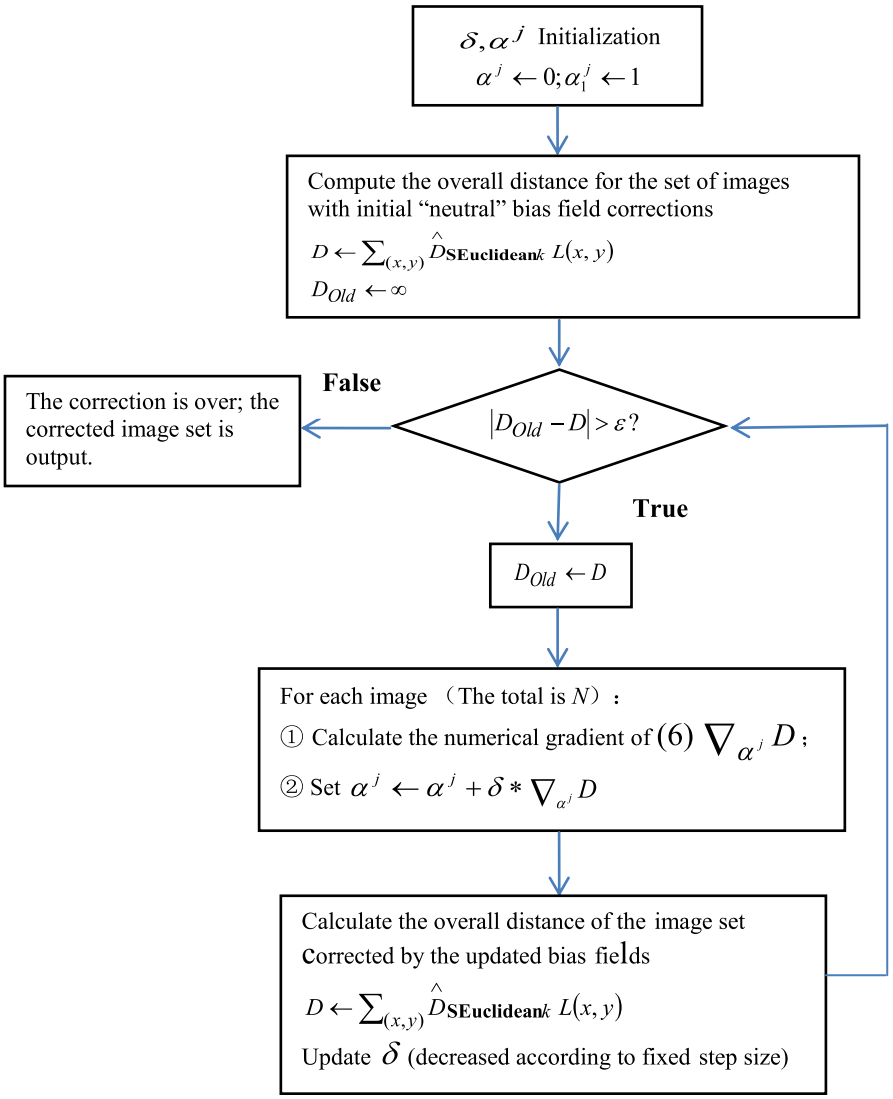

FIGURE 3. Flow chart of the bias field correction algorithm based on square Euclidean distance measure.

set. Therefore, the j-th bias field is calculated as the following [17]:

$$B^j(x,y) = \sum_{k=1}^{25} \alpha_k^j \phi_k(x,y), \tag{7}$$

where ϕ_k represents the k-th two-dimensional sine/cosine base.

Equation (5) or Equation (6) is optimized by the gradient descent method, which simultaneously updates the bias field for each image to reduce the overall distance. In the optimization process, at time step t, the coefficient of each bias field α_k^j for each bias field is updated using the intrinsic image estimation and overall distance estimate from time step $t-1$. After all α^j have been updated, a new set of estimated intrinsic images and overall distances are renewed, and another optimization step is applied until the stop criteria are met. If the gradient step at time step t-1 is δ_{t-1}, then the gradient step at time step t is:

$$\delta_t = \beta * \delta_{t-1}, \quad 0 < \beta < 1. \tag{8}$$

Taking Equation (6) as an example, the algorithm pipeline is shown in Fig. 3

With the right convergence, it is assumed that the distance has been reduced as much as possible by optimizing the bias fields. From the experimental results, by increasing the number of images (N) in the optimization, the likelihood of becoming stuck in local minima is reduced, and the removal effect of bias fields is improved.

The numerical gradient $\nabla_{\alpha^j} D$ of Equation (4), Equation (5) and Equations (6) correspond to the following:

$$\nabla_{\alpha^j} \hat{H}_{Vasicek}\left(Z^1, \ldots, Z^N\right)$$

$$= \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{m}{N}\cdot\frac{1}{Z^{(i+m)}-Z^{(i)}}\cdot\frac{\partial\left(Z^{(i+m)}-Z^{(i)}\right)}{\partial\alpha^j}\right)$$
$$= \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{m}{N}\cdot\frac{1}{Z^{(i+m)}-Z^{(i)}}\cdot\frac{\partial\left(Z^{(i+m)}-Z^{(i)}\right)}{\partial B^j}\cdot\frac{\partial B^j}{\partial\alpha^j}\right)$$
$$= \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{m}{N}\cdot\frac{1}{Z^{(i+m)}-Z^{(i)}}\cdot\frac{\partial\left(Z^{(i+m)}-Z^{(i)}\right)}{\partial B^j}\cdot\frac{\partial\sum_{k=1}^{25}\alpha_k^j\phi_k(x,y)}{\partial\alpha^j}\right). \quad (9)$$

$$\nabla_{\alpha^j}\hat{D}\,Manhat\tan\left(Z^1,\ldots,Z^N\right)$$
$$= \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{N}{m}\cdot\frac{\partial\left(Z^{(i+m)}-Z^{(i)}\right)}{\partial\alpha^j}\right)$$
$$= \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{N}{m}\cdot\frac{\partial\left(Z^{(i+m)}-Z^{(i)}\right)}{\partial B^j}\cdot\frac{\partial B^j}{\partial\alpha^j}\right)$$
$$= \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{N}{m}\cdot\frac{\partial\left(Z^{(i+m)}-Z^{(i)}\right)}{\partial B^j}\cdot\frac{\partial\sum_{k=1}^{25}\alpha_k^j\phi_k(x,y)}{\partial\alpha^j}\right) \quad (10)$$

$$\nabla_{\alpha'}\hat{D}\,\text{sEudiden}k\left(Z^1,\ldots,Z^N\right)$$
$$= \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{N}{m}\cdot 2\left(Z^{(i+m)}-Z^{(i)}\right)\cdot\frac{\partial\left(Z^{(i+m)}-Z^{(i)}\right)}{\partial\alpha^j}\right)$$
$$= \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{N}{m}\cdot 2\left(Z^{(i+m)}-Z^{(i)}\right)\cdot\frac{\partial\left(Z^{(i+m)}-Z^{(i)}\right)}{\partial B^j}\cdot\frac{\partial B^j}{\partial\alpha^j}\right)$$
$$= \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{N}{m}\cdot 2\left(Z^{(i+m)}-Z^{(i)}\right)\cdot\frac{\partial\left(Z^{(i+m)}-Z^{(i)}\right)}{\partial B^j}\cdot\frac{\partial\sum_{k=1}^{25}\alpha_k^j\phi_k(x,y)}{\partial\alpha^j}\right). \quad (11)$$

The measure of the execution time of the algorithm is called the asymptotic time complexity of the algorithm, referred to as the time complexity [24]. Assuming the number of images is N, the time complexity of Equation (10) is $O(q)$, then the time complexity of Equation (9) is $O\left(N^3\right)\cdot O(q)$, and the time complexity of Equation (11) is $O(N)\cdot O(q)$. That is why, in the case of the same number of images, Erik's method [17] runs much longer than our methods.

However, this model also has some shortcomings. For each set of identical images, multiplying each image by the same bias field does not increase the pixel-stack distance (since all images will still be the same). More generally, when the images are approximately equal, their pixel-stack distances are minimally affected by a "common" bias field, i.e., the bias field that appears in all images. Therefore, the model cannot, in general, eliminate all bias fields from a set of images but only sets all of the bias fields as equivalent. Thus any constant bias field that remains in all images after correction is called the residual bias field.

If the residual bias field tends to contain the components of each base, its value approximates the average of the components across images. For example, if half of the contaminated images have a positive value for a particular component's coefficient and the other contaminate images have a negative coefficient for that component, the residual bias field tends to have a coefficient of the component that is near zero. However, if the same type of bias field component appears in most of the images, the algorithm does not remove it. In this case, a within-image method can be used to further remove the bias, but there is a risk of removing components that actually represent smooth gradations in the anatomy.

IV. QUANTITATIVE STANDARD FOR BIAS REMOVAL

A. IMAGE INFORMATION ENTROPY

Entropy is a measure of information unpredictability, which can effectively evaluate the degrees of intensity inhomogeneity. Smaller entropy indicates a homogeneous image.

For image I, the gray scale range is $1 \sim M$, and P is a random variable within gray set $\boldsymbol{G}$ $\{1, 2, \ldots, M\}$; its probability distribution is $\{p(i) = p\{P = i\}, i \in \mathbf{G}\}$, where $p(i)$ denotes the probability of occurrence of pixels with gray value i in image I, i.e., the value of the corresponding gray level i in the gray histogram. The entropy of image I is defined as the entropy of the random variable P:

$$H(P) = -\sum_{i\in\mathbf{G}} p(i)\log(p(i)) = H(I). \quad (12)$$

$H(P)$ is a unimodal convex function distribution. When the event obeys a uniform distribution, the entropy value at this time is the largest, that is, $p(i) = 1/M, i = 1, 2, \ldots, M, H(P)_{\max} = \log M$.

The image entropy is a statistical feature that reflects the amount of average information in an image and is a measure of the uncertainty of a random variable. When the entropy is greater, the uncertainty of the event occurrence is greater, and hence, the amount of information that can be obtained from random variables is greater [25]. For a real undisturbed image, the amount of information (i.e., entropy) is taken as a fixed value H_1, and the entropy of the contaminated image of the bias field is H_2. Due to the ambiguity of the bias field, the probability of gray level occurrence of the contaminated image tends toward an average distribution, resulting in an increase in the uncertainty of the variable P, that is $H_1 < H_2$. Therefore, image entropy can be used as a quantitative criterion for evaluating the image sharpness.

B. SSIM

The similarity measurement is performed on the intrinsic image and the in-homogeneity-corrected image, and the quality of bias field removal is judged according to the image similarity. Common image similarity comparison indicators are peak signal to noise ratio (PSNR) and structural similarity index (SSIM). The PSNR is based on the error between the corresponding pixel points, that is, the image quality evaluation based on error sensitivity, and does not take into account the visual characteristics of the human eye (the human eye is more sensitive to contrast differences of the lower spatial frequencies, and the contrast difference in brightness is more sensitive than the contrast in chromaticity; the perception result of the human eye in an area is affected by the surrounding neighborhood), such that the evaluation result is often inconsistent with the subjective feeling of the person. SSIM is a structural similarity index that measures the similarity between two images. Compared with the PSNR, structural similarity is more consistent with human visual characteristics in evaluating image quality [26]. SSIM is superior to the PSNR in image denoising and image similarity evaluation. Therefore, in this paper, SSIM is chosen as the similarity measure of the intrinsic image and the corrected image in Section V.A.

SSIM is also a full-reference image quality evaluation index that measures the image similarity from three aspects of brightness, contrast and structure [26]–[28]. The SSIM value range is [0, 1], and a larger value means that the image distortion is smaller, which means that the effect of bias removal is better.

$$SSIM\,(X,Y) = [l\,(X,Y)]^{\alpha}\,[c\,(X,Y)]^{\beta}\,[s\,(X,Y)]^{\gamma} \quad \alpha > 0, \beta > 0, and, \gamma > 0, \tag{13}$$

where

$$l\,(X,Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1}, \quad c\,(X,Y) = \frac{2\sigma_{XY} + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2},$$

and $s\,(X,Y) = \frac{\sigma_{XY}+C_3}{\sigma_X\sigma_Y+C_3}$ indicate the brightness similarity, contrast similarity, and structural similarity of the two images, respectively; μ_X and, μ_Y indicate the means of the images X and Y, respectively; σ_X and, σ_Y indicate the variances of the images X and Y, respectively; and σ_{XY} indicates the covariance between images X and Y:

$$\mu_X = \frac{1}{V \times W}\sum_{i=1}^{V}\sum_{j=1}^{W} X\,(i,j)$$

$$\sigma_X^2 = \frac{1}{V \times W - 1}\sum_{i=1}^{V}\sum_{j=1}^{W} (X\,(i,j) - \mu_X)^2,$$

$$\sigma_{XY} = \frac{1}{V \times W - 1}\sum_{i=1}^{V}\sum_{j=1}^{W} ((X\,(i,j) - \mu_X)\,(Y\,(i,j) - \mu_Y)).$$

C_1, C_2, C_3 are constants, which are used to avoid a denominator of 0, and they are usually taken as $C_1 = (K1{*}R)^2$, $C_2 = (K2 * R)^2$, $C_3 = \frac{C_2}{2}$ respectively, and generally $K1 = 0.01, K2 = 0.03, R = 255$.

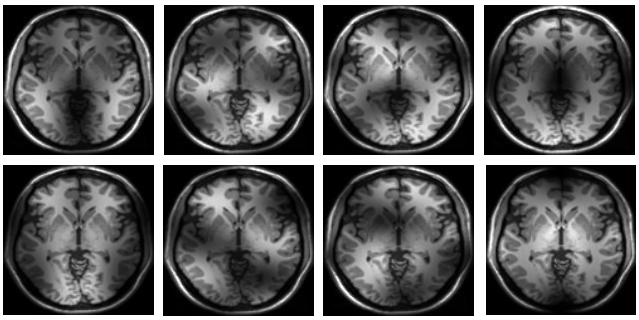

FIGURE 4. Eight images contaminated by bias.

In actual application, for simplicity, the parameters, α, β and γ are generally set to 1, then SSIM is simplified to the following:

$$SSIM\,(X,Y) = \frac{(2\mu_X\mu_Y + C_1)\,(2\sigma_{XY} + C_2)}{\left(\mu_X^2 + \mu_Y^2 + C_1\right)\left(\sigma_X^2 + \sigma_Y^2 + C_2\right)}. \tag{14}$$

C. COEFFICIENT OF JOINT VARIATION

Coefficient of variation (CV) can only calculate the uniformity of a kind of tissue, while coefficient of joint variation (CJV) uses the gray information of two kinds of tissues, can simultaneously evaluate the uniformity of the two tissues, and can better reflect the gray level uniformity of the whole image, and only need to know the rough, conservative segmentation results [29]–[31].

The coefficient of joint variation (CJV) between the two tissues A and B is defined as follows [1]:

$$CJV\,(A,B) = \frac{\sigma\,(A) + \sigma\,(B)}{|\mu\,(A) + \mu(B)|}, \tag{15}$$

where μ and σ represent the mean and standard deviation of tissue class respectively. It can be seen from the definition that the more uniform the gray level of the image, the smaller the CJV value obtained by the two tissues. Conversely, the smaller the CJV value of the two tissues, the better the image gray uniformity.

V. ANALYSIS OF EXPERIMENTAL RESULTS

To verify the proposed algorithm in this paper using nonparametric maximum likelihood to jointly eliminate multiple MRI image bias, we compared our methods with Erik's method [17] and N4 method [12] on a synthetic dataset and a real MRI dataset. The experimental platform is an Intel(R) Core(TM) i5-2400 CPU @3.10GHz, and 4G memory, and the operating system is Windows 7.

A. CORRECTION OF SIMULATED MRI IMAGES

A bias-free, noise-free synthetic brain image was obtained from the Brain Web Simulated Brain Database [20] as an intrinsic image, as shown in Fig. 1(a). The size of the image is 120 × 120. Using *k (k >= 5)* copies of a single intrinsic

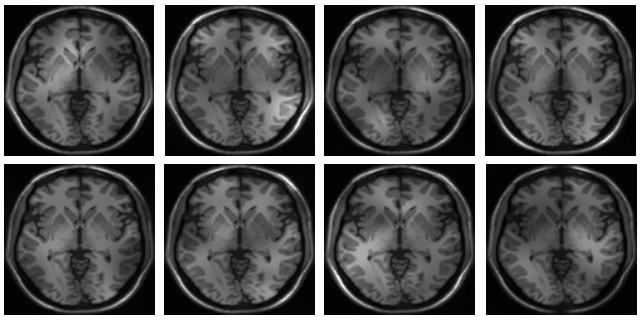

FIGURE 5. Images corrected by Erik's method [18].

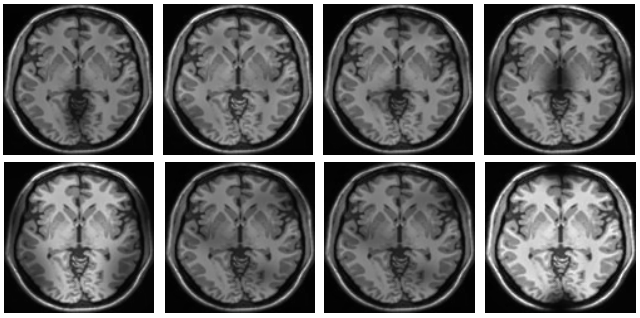

FIGURE 6. Images corrected by N4 method [12].

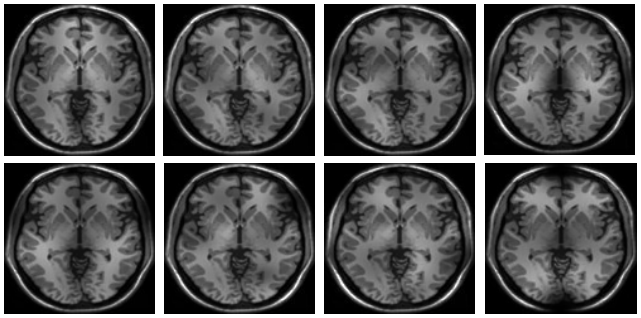

FIGURE 7. Images corrected by our method (MMD).

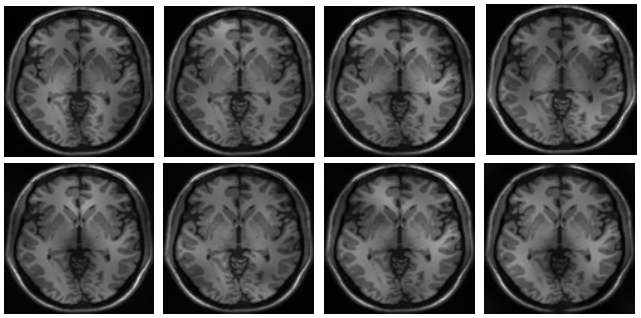

FIGURE 8. Images corrected by our method (MSED).

image, we added different simulated bias fields to each one. We refer to the method of minimizing the Manhattan distance (1-norm) function measure as the MMD algorithm and the

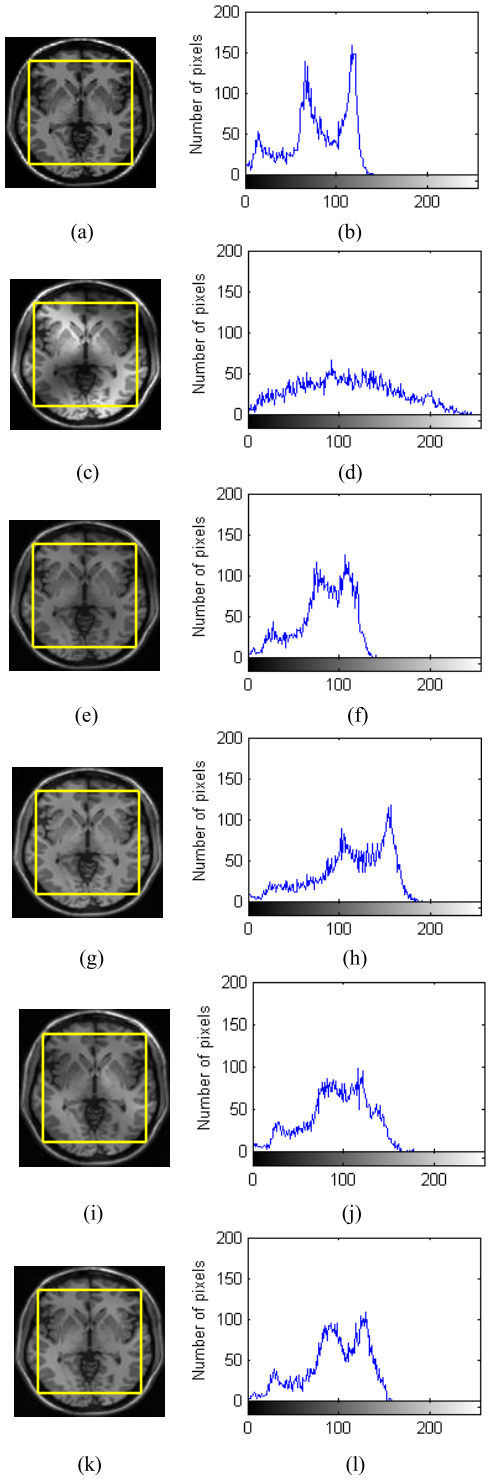

FIGURE 9. Influence of intensity inhomogeneity on intensity distribution. (a) An intrinsic image. (b) The corresponding gray level histogram of the yellow square in (a). (c) The contaminated image. (d) The corresponding gray level histogram of the yellow square in (c). (e) The corrected image by Erik's method [18]. (f) The corresponding gray level histogram of the yellow square in (e). (g) The corrected image by N4 method [13]. (h) The corresponding gray level histogram of the yellow square in (g). (i) The corrected image by our method (MMD). (j) The corresponding gray level histogram of the yellow square in (i). (k) The corrected image by our method (MSED). (l) The corresponding gray level histogram of the yellow square in (k).

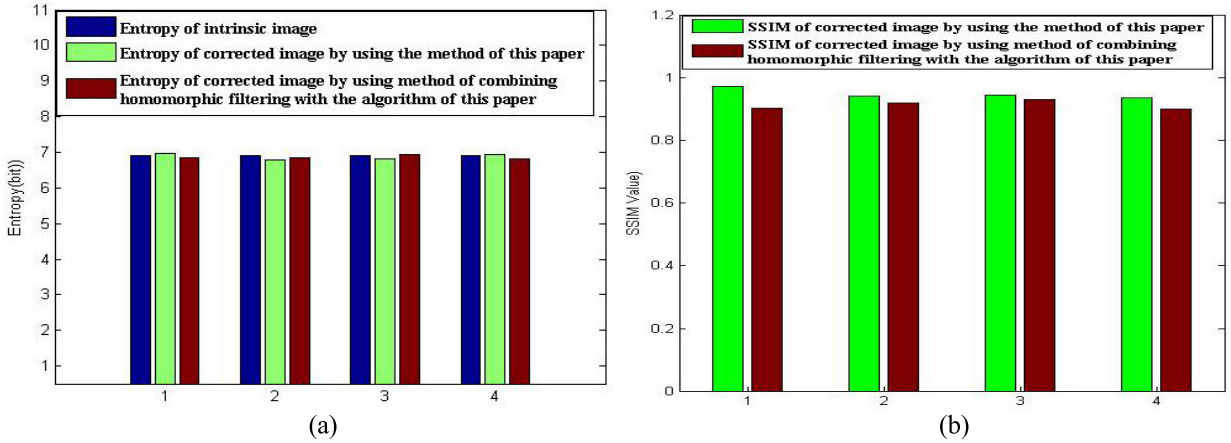

FIGURE 10. Comparison of entropy and SSIM values between the corrected image and the intrinsic image. (a) Entropy of the intrinsic image and the corrected image. (b) The SSIM value of the intrinsic image and the corrected image.

TABLE 1. The quantitative evaluation of bias removal of 8 simulated MRI images.

	Intrinsic images	Contaminated images	Bias field corrected data			
			Erik's [17]	N4 [12]	Proposed (MMD)	Proposed (MSED)
Overall entropy	**55.1516**	57.1383	55.1795	55.4716	55.5519	**55.1735**
Average SSIM with intrinsic image		0.8436	0.9159	0.9288	0.9196	0.9146
Average CJV	0.4878	0.9085	0.6701	0.8176	0.6663	**0.6531**
Running time (s)			35.6	13.8	2.2	11.6

method of minimizing the squared Euclidean distance (2-norm) function measure as the MSED algorithm in this paper.

The experimental results are shown as follows:

From the experimental results, these four methods can better remove the bias of the images. Only the N4 method is to remove the bias of each image individually, and the other three methods are to simultaneously remove the bias of the eight images.

Histogram can well reflect the difference between original images and bias corrected images. We test these four methods on the third contaminated image. The intrinsic image, contaminated image, and bias corrected images are shown in the first column and corresponding gray level histograms are shown in the second column respectively in Fig.9.

The contaminated image (Fig. 9(c)) exhibits obvious intensity inhomogeneities, and its histogram in Fig. 9(d) do not have well-separated peaks due to intensity inhomogeneities. In contrast, it can be seen that intensities within each tissue become quite homogeneous in the bias corrected images, and the gray level histograms of bias corrected images in Fig. 9(f) and Fig. 9(l) appears five well-separated and well-defined peaks. The form of Fig. 9(l) is closest to the gray level histogram of the intrinsic image in Fig. 9(b). The gray level histogram in Fig. 9(h) only appears three well-separated and the trough value is not obvious.

Currently, no gold standard exists for the quality of MRI bias field removal. Although entropy can be used as a evaluation parameter for correction [32], [33], there is a certain error in judging the effect of bias removal only by relying on the reduction of image entropy.

In this paper, the 8 contaminated images are preprocessed by homomorphic filtering [3], and the MSED method is used to remove the bias field among these images (called method G4 in this paper). The first four images of the corrected images and the first four images in Fig. 8 were selected, the entropy of each image and the SSIM value of each image

TABLE 2. The quantitative evaluation of bias removal of 12 simulated MRI images.

	Intrinsic images	Contaminated images	Bias field corrected data			
			Erik's [17]	N4[12]	Proposed (MMD)	Proposed (MSED)
Overall entropy	**82.7274**	85.6458	82.8156	83.4602	82.4763	**82.6984**
Average SSIM with intrinsic image		0.8288	0.9325	0.9193	0.9198	**0.9512**
Average CJV	0.4878	0.8743	0.6252	0.8261	0.6598	**0.5952**
Running time (s)			74.6	20.6	**3.2**	20.8

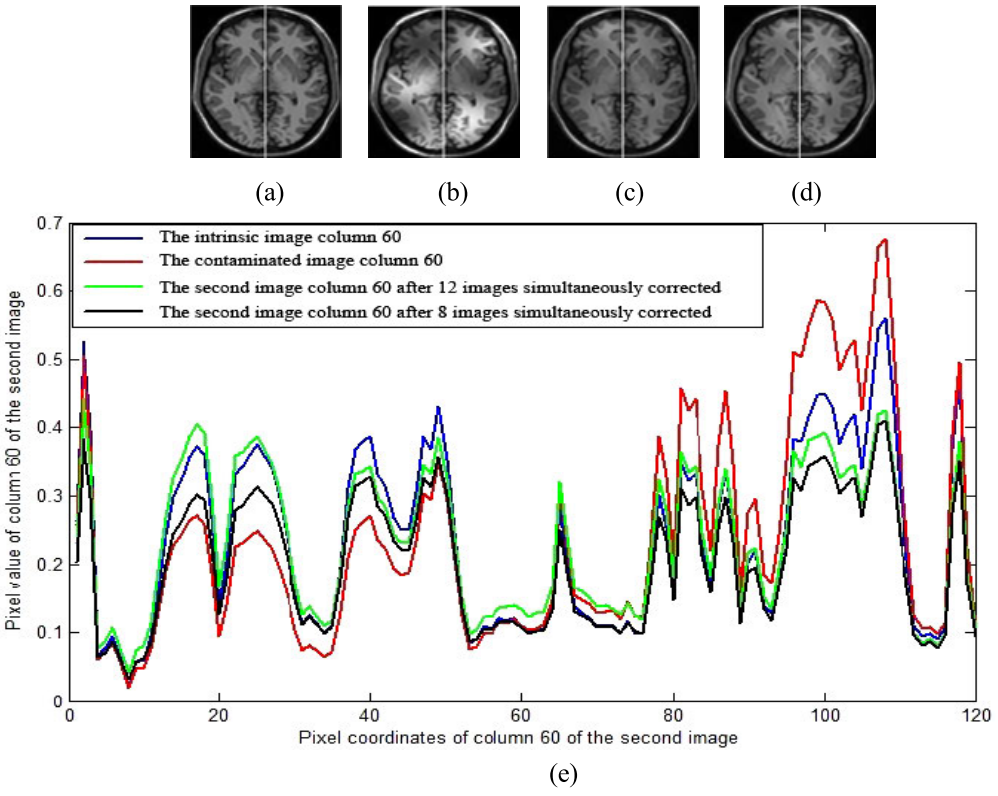

FIGURE 11. Comparison of a simulated human brain MR image contaminated by the bias field and corresponding corrected images. (a) Intrinsic image. (b) Contaminated image. (c) The result of (b) image after bias correction of 8 images simultaneously. (d) The result of (b) image after bias correction of 12 images simultaneously. (e) Distribution of pixel values of column 60 of (a), (b), (c), (d).

and Fig. 1(a) were calculated. The experimental results are shown in Fig. 10.

From Fig. 10(a), the brown column is closest to the blue column, that is, after the bias field is removed by the G4 method, the entropy of the corrected image is closest to the entropy of the intrinsic image. However, from Fig. 10(b), the green column is higher than the brown column, which shows that the SSIM value of the MSED method is larger than that of the method of G4, that is, each image of Fig. 8 is closest to the intrinsic image. Based on this result, we compared our methods with Erik's [17] method and N4 [12] method in terms of overall entropy, average SSIM values and average CJV values of the corrected images, and algorithm running time, which are listed in Table 1 and Table 2.

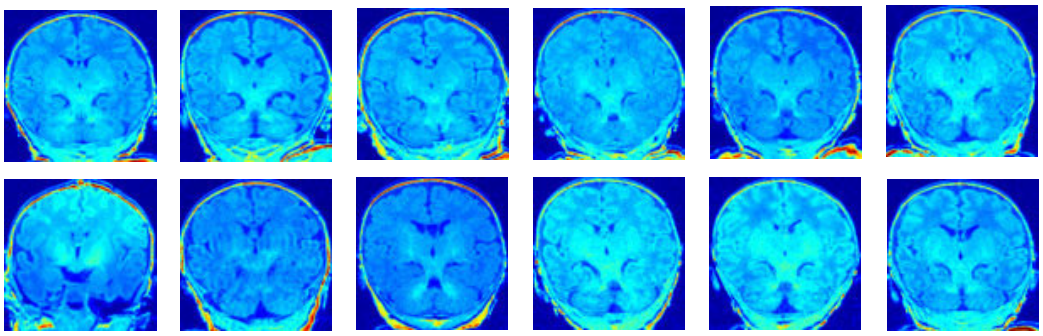

FIGURE 12. Brain MR color contaminated images of 12 different infants.

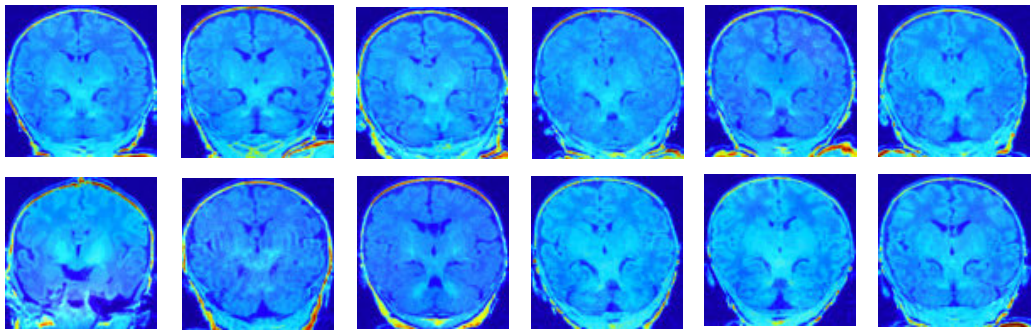

FIGURE 13. Color images corrected by Erik's method [18].

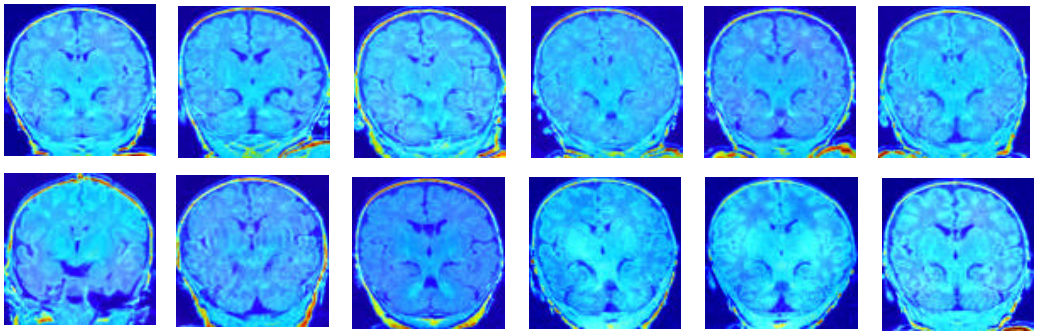

FIGURE 14. Color images corrected by N4 method [13].

From Table 1, it can be observed that the overall entropy of the 8 intrinsic images is 55.1516, and that of the 8 contaminated images is 57.1383, the average SSIM value and average CJV value of contaminated images are 0.8436 and 0.9085, respectively. After bias correction, the overall entropy of our MSED method is 55.1735, which is the closest to the overall entropy of the intrinsic images. As previously stated, the smaller the CJV is, the better the image uniformity is. Due to the influence of the bias field, the CJV of the contaminated image is higher than the CJV of the intrinsic image. The CJV of our MSED method is 0.6531, which is the smallest of the four methods, which means that the gray uniformity of gray matter and white matter is the best after the bias removal by this method. The average SSIM of N4 [12] method is 0.9288, which is the highest of the four methods. However, as the number of images increases, the advantages of our algorithm will become more and more obvious.

In Table 2, when the number of images increases to 12, we can see that the error between the overall entropy of the images corrected by our MSED method and that of the intrinsic images is 0.029, which is the smallest of the four methods, and the corresponding average SSIM value and average CJV value are the best of the four methods, that implies that the images corrected by our MSED method are closest to the intrinsic images. The average SSIM value of the N4 [12] method is not affected by the number of images, and is basically maintained at around 0.92. Although our MMD method is not the best, it takes the least time. It takes only 3.2 seconds for 12 images simultaneously to remove the bias.

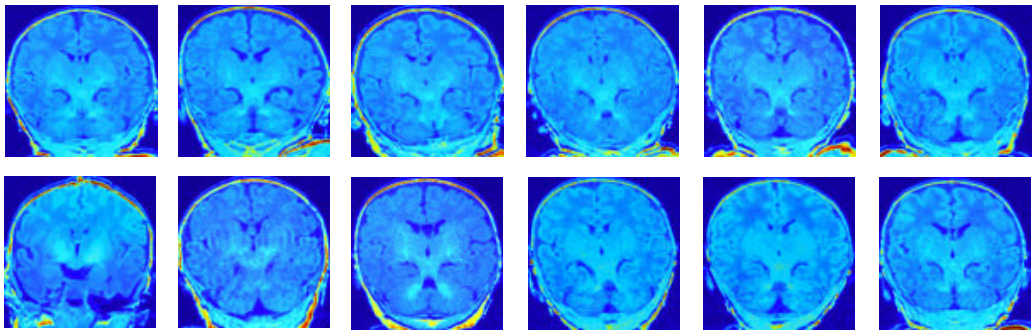

FIGURE 15. Color images corrected by our method (MMD).

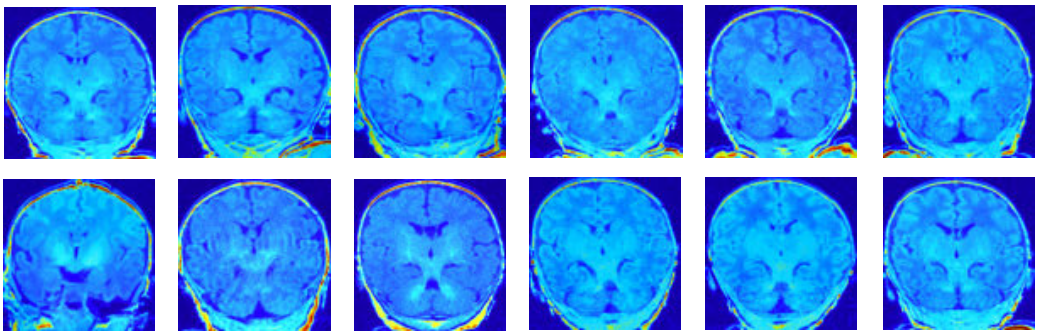

FIGURE 16. Color images corrected by our method (MSED).

TABLE 3. The quantitative evaluation of bias removal of 12 real MRI images.

	Contaminated images	Bias field corrected data			
		Erik's [17]	N4 [12]	Proposed (MMD)	Proposed (MSED)
Overall entropy	-1.3308e+004	-1.5136e+004	-1.1825e+004	-1.5373e+004	-1.4868e+004
Average CJV	0.7171	0.6975	0.7593	0.6914	**0.6872**
Running time (s)		28.8	9.2	**2.3**	5.5

The larger the number of images in this model is, the better the bias removal effect is. For each corresponding method in Table 1 and Table 2, the average SSIM values and the average CJV value of the 12 images are better than those of the 8 images except for the N4 [12] method. The larger the SSIM value is, the higher the similarity with the intrinsic image is, and the better the bias removal effect is. At the same time, the smaller the value of CJV is, the better image uniformity is.

In the following experiment, our MSED method is used to simultaneously remove the 8 contaminated images and 12 contaminated images, and the size of each image is 120×120. By respectively taking the pixel values of the 60th column (indicated by the white line in Fig. 11) of the intrinsic image, the second contaminated image and the second image in the two groups of corrected images are compared, as shown in Fig. 11(e).

The blue line, red line, black line, and green line are the distributions of pixel values in the 60th columns of Fig. 11(a), Fig. 11(b), Fig. 11(c), and Fig. 11(d), respectively. The comparison shows that the fitting effect of the green line and the blue line is better than that of the black line and the blue line, that is, Fig. 11(d) is closer to the intrinsic image. This result also demonstrates that the greater the number of images is, the better the bias removal effect of this model is.

B. CORRECTION OF REAL MRI IMAGES

In this experiment, the image set was taken from http://people.cs.umass.edu/~elm/congealing/, and 12 different infant brain images were selected (shown in Fig. 12).

The bias field can be seen in many images. The most striking is a ''ramp-like'' bias field in the first image of the second row (The top of the brain is too bright, and the bottom is too dark.). Because the white matter of the brain is not fully developed during these infant scans, it is difficult to classify the tissue into a fixed number of categories, as is usually performed for adult brain images, and therefore, these images are not amenable to methods based on specific tissue models developed for adults (e.g., [13]).

The bias fields of the 12 different infant brain images were removed by Erik's [17] method, N4 [12] method, our MMD method, and our MSED method. The experimental results are shown in Fig. 13, Fig. 14, Fig. 15, and Fig. 16 respectively.

Carefully comparing the central brain regions in Fig. 13, Fig. 14, Fig. 15, and Fig. 16, we can see that the butterfly shaped region in the middle of the brain, which represents developing white matter, has been suppressed in Fig. 14. This is most likely because the N4 [12] algorithm strives to make the image more uniform by removing the bright part in the middle of the image. However, the other three algorithms which compares pixels across images, do not suppress these real structures since they occur across images. Hence, operations across images can sometimes produce superior results.

In Fig. 17, we present the histograms for the seventh original image and the corresponding bias corrected images. It shows that the intensities of the bias corrected images are more concentrated on one or more constant values compared with the original image.

According to the histogram distribution, the pixels of Fig. 17(f) are mainly concentrated in the peak position, in combination with Fig. 17(e), the distinction between gray matter and white matter is not obvious. Although there is a good distinction between gray matter and white matter in Fig. 17(c), Fig. 17(j) and Fig. 17(i), from the corresponding histogram distribution, we can see that the pixel values in Fig. 17(d) fluctuate greatly at the peak position, while in Fig. 17(h) and Fig. 17(j), the pixel values fluctuate less at the peak position, which indicates that intensities of each tissue becomes more homogeneous in the bias corrected images by our methods.

Since there are no intrinsic images, we are unable to calculate the SSIM value between the corrected image and the intrinsic image. In this section, we compared our methods with Erik's [17] method and N4 [12] method in terms of overall entropy, average CJV values of the corrected images, and algorithm running time, which are listed in Table 3. The overall entropy of the images is calculated by using the entropy estimator $\hat{H}_{Vasicek}$ (Formula (4)).

It can be seen from Table 3 that the overall entropy of the N4 [12] method does not decrease, but is larger than the overall entropy of the original images, while the overall entropy of the other three methods decreases in varying degrees compared to that of the original images. Similarly, the average CJV value of N4 [12] method is higher than that of the original image, which is due to the unclear boundary

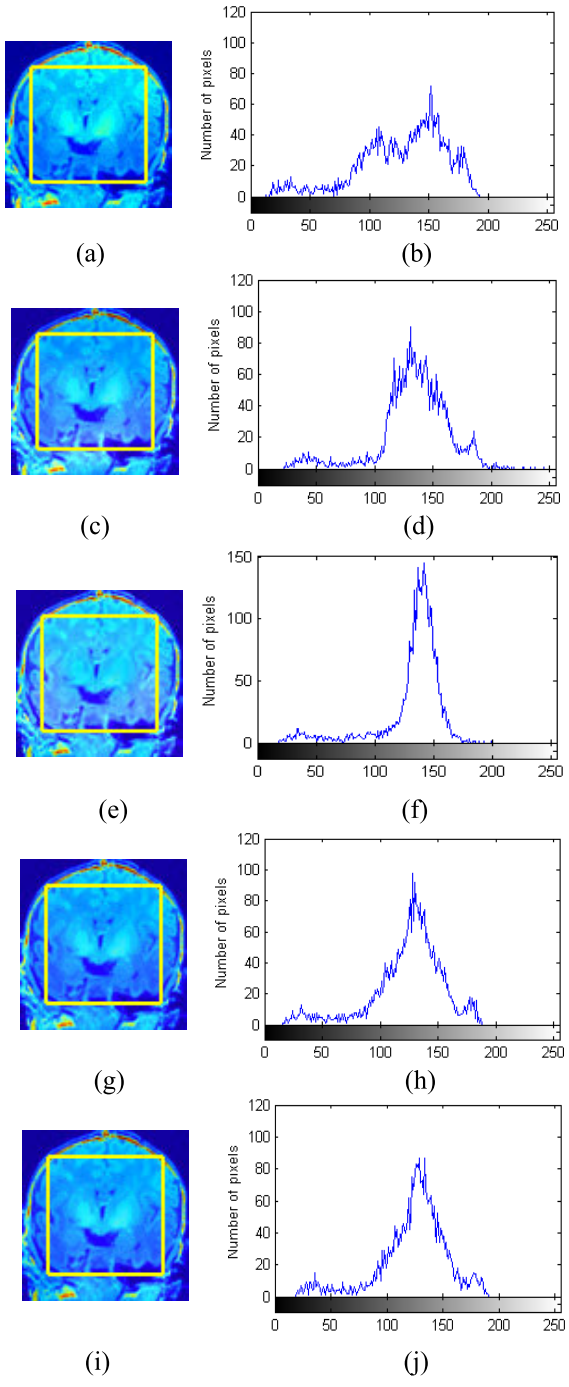

FIGURE 17. Influence of intensity inhomogeneity on intensity distribution. (a) The seventh original image. (b) The corresponding gray level histogram of the yellow square in (a). (c) The corrected image by Erik's method [18]. (d) The corresponding gray level histogram of the yellow square in (c). (e) The corrected image by N4 method [13]. (f) The corresponding gray level histogram of the yellow square in (e). (g) The corrected image by our method (MMD). (h) The corresponding gray level histogram of the yellow square in (g). (i) The corrected image by our method (MSED). (j) The corresponding gray level histogram of the yellow square in (i).

between gray matter and white matter in the bias corrected images by N4 [12] method. This is consistent with our previous conclusion from the histogram. The average CJV value of our MSED method is the lowest of the four methods. Although the quantitative evaluation of our MMD method is not the best, the speed is the fastest, it takes only 2.3 seconds for 12 images simultaneously to remove the bias, while Erik's [17] method takes 28.8 seconds.

From Section V, we can see whether on the synthetic MRI images or on the real MRI images, the correction effect of our MSED algorithm is the best, and the running speed is much faster than the Erik's [17] method. In Section V.A., the bias fields of 12 synthetic images with 120×120 specifications were removed simultaneously, the running speed of our MESD method was approximately 3.5 times faster than that of the Erik's [17] method. In Section V.B., the bias fields of 12 real images with 90×90 specifications were removed simultaneously, our MESD method was approximately 5.2 times faster than the Erik's [17] method. As the number of images increases, the speed advantage of the MSED algorithm becomes more obvious.

VI. CONCLUSION

The problem of bias field exists in an MR image because it is contaminated by the multiplicative bias field. Although researchers have proposed many correction methods, no common method exists to date, and no objective standard is available for judging whether the correction is successful.

To settle the problem described above, this paper proposes a method for removing the MRI bias by minimizing the distance measure across images and uses three objective criteria of overall entropy, average SSIM value, and average CJV value to assess the effect of bias correction on synthetic MRI images and real MRI images. The algorithm uses two types of information, within and across images to remove the bias of a group of images simultaneously. The greater the number of images is, the lower the probability of convergence being limited to local minima is, and the better the effect of bias removal is.

For the ''simple'' bias correction problem, this method may be excessive, but for difficult problems in bias correction, such as the case in which the bias field is difficult to separate from the underlying tissue, as discussed in [34], this method works well, and the organizational structure is not suppressed. At present, most of the bias correction algorithms mainly use 2D images as research objects, which are validated on 2D or 3D data. In the future work, we will consider: (1) bias correction of 3D MRI images; (2) to combine the spatial transformation method to simultaneously perform group registration and bias removal on the images in this model.

REFERENCES

[1] B. Likar, M. A. Viergever, and F. Pernus, ''Retrospective correction of MR intensity inhomogeneity by information minimization,'' *IEEE Trans. Med. Imag.*, vol. 20, no. 12, pp. 1398–1410, Dec. 2001.

[2] J. C. Rajapakse and F. Kruggel, ''Segmentation of MR images with intensity inhomogeneities,'' *Image Vis. Comput.*, vol. 16, no. 3, pp. 165–180, 1998.

[3] Y. Li, Z. Wang, and Y. S. Zhu, ''Inhomogeneity correction of magnetic resonance images,'' *J. Shanghai Jiaotong Univ. Sci. Technol., Nature Sci.*, vol. 37, no. 9, pp. 1452–1455, 2003.

[4] B. R. Condon, J. Patterson, D. Wyper, A. Jenkins, and D. M. Hadley, ''Image non-uniformity in magnetic resonance imaging: Its magnitude and methods for its correction,'' *Brit. J. Radiol.*, vol. 60, no. 1, pp. 83–87, 1987.

[5] A. Simmons, P. S. Tofts, G. J. Barker, and S. R. Arridge, ''Sources of intensity nonuniformity in spin echo images at 1.5 T,'' *Magn. Reson. Med.*, vol. 32, no. 11, pp. 121–128, 1994.

[6] H. Wu, ''Research on correction method of intensity inhomogeneity in magnetic resonance image,'' M.S. thesis, Dept. Elect. Sci. Technol., Univ. Sci. Technol. China, Hefei, China, May 2008.

[7] Z. Hou, ''A review on MR image intensity inhomogeneity correction,'' *Int. J. Biomed. Imag.*, vol. 2006, Feb. 2006, Art. no. 49515.

[8] U. Vovk, F. Pernus, and B. Likar, ''A review of methods for correction of intensity inhomogeneity in MRI,'' *IEEE Trans. Med. Imag.*, vol. 26, no. 3, pp. 405–421, Mar. 2007.

[9] R. Guillemaud, ''Uniformity correction with homomorphic filtering on region of interest,'' in *Proc. IEEE Int. Conf. Image Process.*, Oct. 1998, pp. 872–875.

[10] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, ''A nonparametric method for automatic correction of intensity nonuniformity in MRI data,'' *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 87–97, Feb. 1998.

[11] M. Styner, C. Brechbuhler, G. Szckely, and G. Gerig, ''Parametric estimate of intensity inhomogeneities applied to MRI,'' *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 153–165, Mar. 2000.

[12] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, ''N4ITK: Improved N3 bias correction,'' *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1310–1320, Jun. 2010.

[13] W. M. Wells, III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, ''Adaptive segmentation of MRI data,'' *IEEE Trans. Med. Imag.*, vol. 15, no. 8, pp. 429–442, Aug. 1996.

[14] P. Viola and W. M. Wells, III, ''Alignment by maximization of mutual information,'' Ph.D. dissertation, Dept. Radiol., Massachusetts Inst. Technol., Cambridge, MA, USA, 1995.

[15] H. Liu, S. Liu, D. Guo, Y. Zheng, P. Tang, and G. Dan, ''Original intensity preserved inhomogeneity correction and segmentation for liver magnetic resonance imaging,'' *Biomed. Signal Process. Control*, vol. 47, pp. 231–239, Jan. 2019.

[16] M. M. George, S. Kalaivani, and M. S. Sudhakar, ''A non-iterative multi-scale approach for intensity inhomogeneity correction in MRI,'' *Magn. Reson. Imag.*, vol. 42, pp. 43–59, Oct. 2017.

[17] E. G. Learned-Miller and P. Ahammad, ''Joint MRI bias removal using entropy minimization across images,'' in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 17, 2005, pp. 761–768.

[18] E. G. Learned-Miller and V. Jain, ''Many heads are better than one: Jointly removing bias from multiple MRIs using nonparametric maximum likelihood,'' in *Information Processing in Medical Imaging*. Berlin, Germany: Springer, 2005, pp. 615–626.

[19] E. G. Learned-Miller, ''Data driven image models through continuous joint alignment,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 2, pp. 236–250, Feb. 2005.

[20] A. C. Evans. *Brain Web: Simulated Brain Database [DB/OL]*. Accessed: Oct. 2018. [Online]. Available: http://brainweb.bic.mni.mcgill.ca/brainweb/selection_normal.html

[21] P. A. Viola and W. M. Wells, III, ''Alignment by maximization of mutual information,'' in *Proc. IEEE Int. Conf. Comput. Vis.*, Jun. 1995, pp. 16–23.

[22] O. Vasicek, ''A test for normality based on sample entropy,'' *J. Roy. Stat. Soc., B Methodol.*, vol. 31, no. 1, pp. 632–636, 1976.

[23] E. G. Learned-Miller and J. W. Fisher, III, ''ICA using spacings estimates of entropy,'' *J. Mach. Learn. Res.*, vol. 4, pp. 1271–1295, Jan. 2003.

[24] D. Jia and X. Zhang, ''Research on time complexity measure method based on analysis method,'' *J. Liaoning Tech. Univ. (Natural Sci. Ed.)*, vol. 35, no. 4, pp. 231–233, 2015.

[25] J. W. Liu, C. F. Li, and H. Wu, ''An improved entropy method for correction of intensity inhomogeneity in MRI data,'' *Chin. J. Biomed. Eng.*, vol. 27, no. 6, pp. 867–871, Dec. 2008.

[26] Y.-B. Tong, Q.-S. Zhang, and Y.-P. Qi, ''Image quality assessing by combining PSNR with SSIM,'' *J. Image Graph.*, vol. 11, no. 12, pp. 1758–1763, 2006.

[27] H. Liu, X. Bi, G. Lu, W. Wang, J. Yan, and Z. Zhang, ''Screen window propagating for image inpainting,'' *IEEE Access*, vol. 10, no. 6, pp. 61761–61772, 2008.

[28] X. Pei, T. Dong, and Y. Guan, "Super-resolution of face images using weighted elastic net constrained sparse representation," *IEEE Access*, vol. 7, no. 5, pp. 55180–55190, 2019.
[29] Z. Y. Chua, W. Zheng, and M. W. L. Chee, "Evaluation of performance metrics for bias field correction in MR brain images," *J. Magn. Reson. Imag.*, vol. 29, no. 6, pp. 1271–1279, 2009.
[30] C. Li, J. C. Gore, and C. Davatzikos, "Multiplicative intrinsic component optimization (MICO) for MRI bias field estimation and tissue segmentation," *Magn. Reson. Imag.*, vol. 32, no. 7, pp. 913–923, Sep. 2014.
[31] M. M. George and S. Kalaivani, "Retrospective correction of intensity inhomogeneity with sparsity constraints in transform-domain: Application to brain MRI," *Magn. Reson. Imag.*, vol. 61, pp. 207–223, Sep. 2019.
[32] O. Salvado, C. Hillenbrand, S. Zhang, and D. L. Wilson, "Method to correct intensity inhomogeneity in MR images for atherosclerosis characterization," *IEEE Trans. Med. Imag.*, vol. 25, no. 5, pp. 539–552, May 2006.
[33] J. S. Suri, R.-F. Chang, G. A. Giraldi, P. S. Rodrigues, J. S. Suri, R.-F. Chang, G. A. Giraldi, and P. S. Rodrigues, "Non-extensive entropy for CAD systems of breast cancer images," in *Proc. Brazilian Symp. Comput. Graph. Image Process.*, Oct. 2006, pp. 121–128.
[34] A. Fan, W. M. Wells, J. W. Fisher, M. Çetin, S. Haker, R. Mulkern, C. Tempany, and A. S. Willsky, "A unified variational approach to denoising and bias correction in MR," in *Information Processing in Medical Imag.*, vol. 18. Berlin, Germany: Springer, 2003, pp. 148–159.

YAN XU received the B.S. degree from Qufu Normal University, Qufu, China, in 2001, and the M.S. degree from Guizhou University, Guizhou, China, in 2006. She is currently pursuing the Ph.D. degree with the Shandong University of Science and Technology, Qingdao, China. She was a Visiting Scholar with China Agricultural University, Beijing, China, in 2009. She is currently a Lecturer with the College of Information Science and Engineering, Linyi University. Her current research interests include image processing, pattern recognition, medical image registration, and deep learning.

SHUNBO HU received the M.S. and Ph.D. degrees from Shandong University, Jinan, in 2003 and 2008, respectively. He is currently a Professor with the Department of Information Science and Engineering, Linyi University, China. He has published more than 50 articles. His current research interests include medical image registration, image segmentation, brain science, machine-learning, and deep learning. He has received two grants from the National Natural Science Foundation of China.

YUYUE DU received the B.S. degree from Shandong University, Jinan, China, in 1982, the M.S. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1991, and the Ph.D. degree in computer application from Tongji University, Shanghai, China, in 2003. He is currently a Professor with the College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao, China. He has published more than 200 articles in domestic and international academic publications, and they are embodied more than 140 times by SCI and EI. His current research interests include formal engineering, Petri nets, real-time systems, web services, and workflows. He is a member of the Professional Committee of Petri Nets of China Computer Federation. He has taken over ten projects supported by the National Natural Science Foundation, the National Key Basic Research Developing Program, and other important and key projects at provincial levels.