# Super-Resolution Reconstruction Using Cross-Scale Self-Similarity in Multi-Slice MRI

Esben Plenge,* Dirk H. J. Poot, Wiro J. Niessen, Erik Meijering

Erasmus MC – University Medical Center Rotterdam
P.O. Box 2040, 3000 CA Rotterdam, The Netherlands
http://www.bigr.nl/

**Abstract.** In MRI, the relatively thick slices of multi-slice acquisitions often hamper visualization and analysis of the underlying anatomy. A group of post-processing techniques referred to as super-resolution reconstruction (SRR) have been developed to address this issue. In this study, we present a novel approach to SRR in MRI, which exploits the high-resolution content usually available in the 2D slices of MRI slice stacks to reconstruct isotropic high-resolution 3D images. Relying on the assumption of local self-similarity of anatomical structures, the method can be applied both to a single slice stack and to the combination of multiple slice stacks that differ in the orientation of their field of view. We evaluate the method quantitatively on synthetic brain MRI and qualitatively on MRI of the lungs. The results show that the method outperforms state-of-the-art MRI super-resolution methods.

**Keywords:** Super-resolution, reconstruction, MRI, self-similarity, dictionaries, cross-scale, brain, lung, inverse problems, regularization.

## 1 Introduction

In magnetic resonance imaging (MRI) sequences requiring long repetition times, conventional 3D imaging usually leads to infeasible scan times, and 2D multi-slice imaging is used instead. However, due to hardware induced limitations on gradient strength, requirements on signal-to-noise ratio (SNR), and other factors, the slices are usually relatively thick compared to the in-plane resolution. Such anisotropy negatively affects visualization and hampers analysis. The isotropy and resolution of images may be improved by super-resolution reconstruction (SRR) methods [1], often divided into 1) methods that base the high-resolution (HR) reconstruction on a single image only, and 2) methods that combine multiple low-resolution (LR) images of the same object acquired under varying fields of views. In both cases, the inverse problem of recovering the HR image is ill-posed, and regularization is applied by exploiting prior knowledge of the HR solution, such as that it must be smooth, piecewise smooth, or sparse.

Recently, a quite different prior has been shown to be very powerful: example-based self-similarity. It is based on the observation that small-scale structures

* Corresponding author

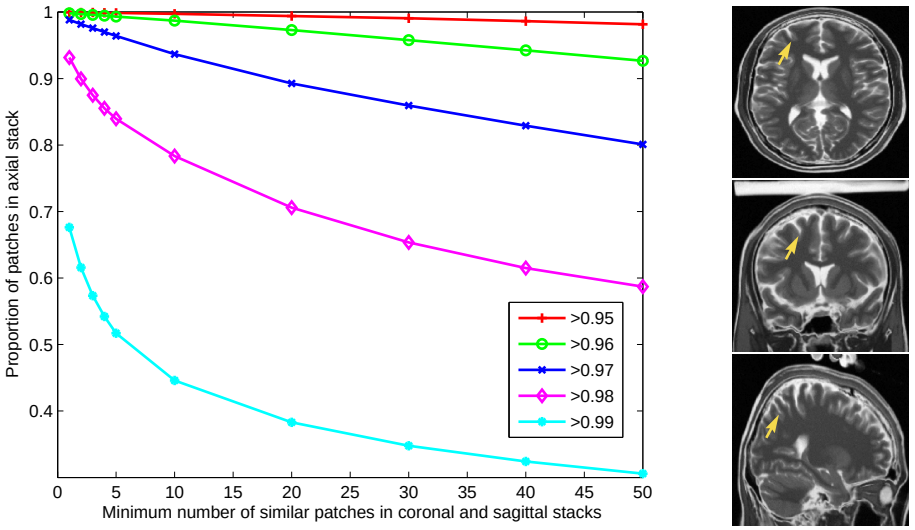



**Fig. 1.** Left: Quantitative check of self-similarity in a T2-weighted brain MRI from BrainWeb [12]. Similarity between the 10% axial 7x7 patches with the highest variance and all coronal and sagittal 7x7 patches was evaluated. For several similarity thresholds, the fraction of axial patches is plotted as function of the number of coronal and sagittal patches exceeding the threshold. Similarity is measured as the inner product between the patches (after vectorization and unit length normalization). Right: Example of local self-similarity in the cortex. The yellow arrows point at a similar cortex structure in the axial (top), coronal (middle), and sagittal (bottom) slices.

tend to repeat themselves throughout an image. This concept has been widely applied for image compression, denoising, inpainting, and SRR [2], and has spawned methodologies such as image hallucination [3], sparse coding using learned dictionaries [4, 5], and non-local means (NLM) [6]. A common feature of these example-based methods is the formulation of a parent-child structure in which the "nearest-neighbor" of an input LR-patch is sought among the LR-parents of HR-child-patches in either a database [3] or in the image itself [7]. In MRI, applications include resolution enhancement of T2-weighted stacks from isotropic HR T1-weighted stacks [8], and upsampling by iteratively applying NLM-denoising to an interpolated version of the image itself [9].

Here we propose a novel approach to SRR in multi-slice MRI based on the concept of *cross-scale* self-similarity [10]. Multi-slice 2D MRI scanning yields two native scales simultaneously: one at in-plane resolution (HR) and one in the slice-selection direction (LR). Local self-similarity of anatomical features occurs both within and across these scales (Fig. 1), which we exploit to achieve SRR. Our method can be applied to both a single image and to multiple images, transcending the methodological division described above. We describe our method and show its potential by comparing it quantitatively and qualitatively with a baseline interpolation scheme and a state-of-the-art SRR algorithm.

## 2 Method

### 2.1 From Single-Scale to Cross-Scale Self-Similarity

Our method is developed from one of the most successful applications of image self-similarity: the non-local means (NLM) method [6, 11]. In NLM, the $i$th pixel in image $\mathbf{u}$, defined over the domain $\Omega$, is represented as a weighted average of pixels from similar patches found elsewhere in the same image:

$$\mathrm{NLM}(\mathbf{u}(i)) = \frac{1}{C(i)} \sum_{j \in \Omega} w(i,j)\mathbf{u}(j) \tag{1}$$

with $w(i,j) = \exp(-\|\mathbf{u}(\mathcal{N}(i)) - \mathbf{u}(\mathcal{N}(j))\|^2_{2,a}/h^2)$ and $C(i) = \sum_{j \in \Omega} w(i,j)$. Here, $\mathcal{N}(k)$ is a square patch of a fixed size centered on pixel $k$, $\|.\|^2_{2,a}$ is the norm weighted by a Gaussian with standard deviation $a$, and $h$ is a parameter that controls the decay of the weights as a function of the norm.

For our purposes, a HR example image $\mathbf{v}$ is used to modify the above NLM expression, leading to cross-scale NLM [10]:

$$\mathrm{csNLM}(\mathbf{u}(i)) = \frac{1}{C(i)} \sum_{j \in \Omega} \omega(i,j)\mathbf{v}(j) \tag{2}$$

with $\omega(i,j) = \exp(-\|\mathbf{u}(\mathcal{N}(\mathcal{D}(i))) - \mathcal{H}_z(\mathbf{v}(\mathcal{N}(j)))\|^2_{2,a}/h^2)$ and $C(i)$ as defined above. Here, $\mathcal{D}$ maps $\mathbf{u}$ onto a HR lattice, and $\mathcal{H}_z$ is a linear model of the imaging process, with degradation $z$. In other words, in csNLM, HR patches $\mathbf{v}(j)$ whose degraded versions are similar to an up-scaled LR patch $\mathbf{u}(i)$ are used to update the HR estimate $\mathrm{csNLM}(\mathbf{u}(i))$.

The method we propose is a special case of csNLM. In the limit, when $h \to 0$, the weight of only one patch (the "nearest neighbor") dominates. The cross-scale super-resolution expression thus becomes:

$$\mathrm{csSR}(\mathbf{u}(i)) = \frac{1}{\omega(i,j)}\omega(i,j)\mathbf{v}(j) = \mathbf{v}(j) \quad \text{with} \quad j = \operatorname*{argmax}_l \omega(i,l). \tag{3}$$

In the experiments we demonstrate that the use of only the nearest neighbor is indeed sufficient to improve resolution and image quality considerably.

### 2.2 MRI Super-Resolution Using Cross-Scale Self-Similarity

From a 2D viewpoint, MR stacks contain both LR images (anisotropic slices containing the slice-selection direction) and HR images (isotropic in-plane slices). Using the knowledge of 2D self-similarity of 3D anatomical structures (Fig. 1), we can apply Eq. (3) as follows: Let $\mathbf{u}$ denote the 2D LR slices of a stack, let $\mathcal{D}$ be an interpolation operator that maps the LR slices onto a 2D HR grid, and $\mathcal{H}_z$ a number of blurring kernels (one of them approximating the point-spread function, PSF, of the upsampled LR slices) applied to the in-plane slices $\mathbf{v}$. We search for the nearest neighbor of $\mathbf{u}(\mathcal{N}(\mathcal{D}(i)))$ over all HR patches $j$, at all

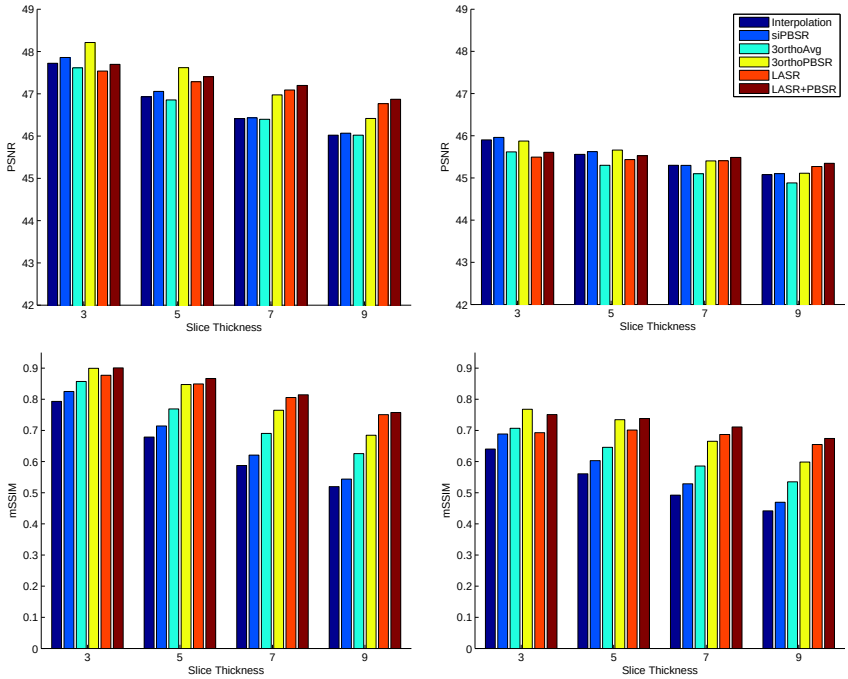



**Fig. 2.** Quantitative evaluations. Top: PSNR of SRR using six different methods. Bottom: mSSIM of SRR by the six methods. Left: Reconstructions from images of the 3% noise group. Right: Reconstructions from images of the 9% noise group.

blurring levels $z$, in $\mathcal{H}_z(\mathbf{v}(\mathcal{N}(j)))$. The reason for using multiple blurring levels is to enrich the model and capture cross-scale self-similarity at scales besides the two native ones (relevant for recursive structures such as vessel-trees). In the experiments we used two levels, empirically chosen to achieve good performance.

**Multiscale Dictionary.** The vectorized versions of all HR patches (or a randomly selected subset thereof) are collected in a matrix. The corresponding vectorized patches of the blurred versions of the HR slices are appended as columns in this matrix, such that each blurring level is represented by a contiguous block of columns. The columns are then normalized and the normalization factors are saved for later rescaling. Adopting the terminology of the sparse coding literature, we shall call this matrix a dictionary.

**Reconstruction Process.** An initial HR estimate is created by interpolating and aligning the LR images onto a 3D HR grid. The slice stack is traversed, LR slice by LR slice. Around each pixel in each slice, a 2D patch is extracted (of size 7×7 pixels in this study). According to the cross-scale self-similarity assumption, the HR version of this LR patch will be well-approximated by the in-plane patches in the dictionary. For each LR patch, the dictionary is searched for its nearest neighbor (NN), defined as the column that has the maximum

inner product with the vectorized and normalized LR patch. The corresponding HR patch is found in the HR block of the dictionary and rescaled to its original intensity. HR patches are inserted into the grid and averaged where they overlap.

## 3 Experiments and Results

### 3.1 Image Data

**Simulated Brain MRI.** Two isotropic HR T2-weighted images (1×1×1 mm voxels) with intensity non-uniformity of 20%, and noise-levels of 3% and 9%, respectively, were downloaded from BrainWeb [12]. To simulate a common case from anatomical neuro-imaging, slice stacks were obtained from each of these HR images by application of a 1D Gaussian PSF in the slice-selection direction. The full-width-half-maximum of the PSF was equal to the downsampling factor applied in the slice-selection direction. By varying this factor as well as the slice-selection direction, three orthogonal slice stacks (axial, coronal, sagittal) were generated for downsampling factors of 3, 5, 7, 9. Thus, a total of eight sets, each containing three orthogonal slice stacks, with resolutions of 1×1×3 mm, 1×1×5 mm, 1×1×7 mm, and 1×1×9 mm, were generated from the two original images with noise-levels of 3% and 9%.

**Lung MRI.** Lung MRI data was acquired on a 1.5 Tesla GE scanner. Axial, coronal, and sagittal slice stacks of a test-subject with the lung in expiratory state were acquired during breath hold. A 2D gradient-recalled steady-state sequence (TR = 2.2 s, TE = 0.75 s, flip angle = 35 degrees, number of averages = 1) was used, having a scan time of 15 seconds per slice stack. The field of view was 400×400 mm, the acquisition matrix was 128×160 pixels, yielding an in-plane resolution of 3.125×2.5 mm. The slices were reconstructed by the scanner to a grid of 256×256 pixels with uniform sizes of 1.56×1.56 mm. 40 slices were acquired per stack with a slice thickness of 8 mm. Lung MRI courtesy of H. Tiddens, P. Ciet, and P. Wielopolski, Dept. of Radiology, Erasmus MC.

### 3.2 Quantitative Evaluation

The proposed method, referred to as patch-based super-resolution (PBSR), was quantitatively evaluated for 1) a single axial slice stack (siPBSR), 2) the combination of three orthogonal stacks (3orthoPBSR), and 3) post-processing of the result of another SRR method (LASR) [13] using three orthogonal stacks (LASR+PBSR). Initial estimates of the HR image were created from the available data: in the case of a single stack, by cubic interpolation, and in the case of three stacks, by either the average of the interpolated and aligned stacks (3orthAvg) or by the LASR method. Performance was quantified by peak-signal-to-noise ratio (PSNR) and mean structural similarity (mSSIM) [14]. To compute these measures, a noise-free HR T2-weighted image (1×1×1 mm voxels) from BrainWeb was used. The results are shown in Fig. 2. As expected, lower noise

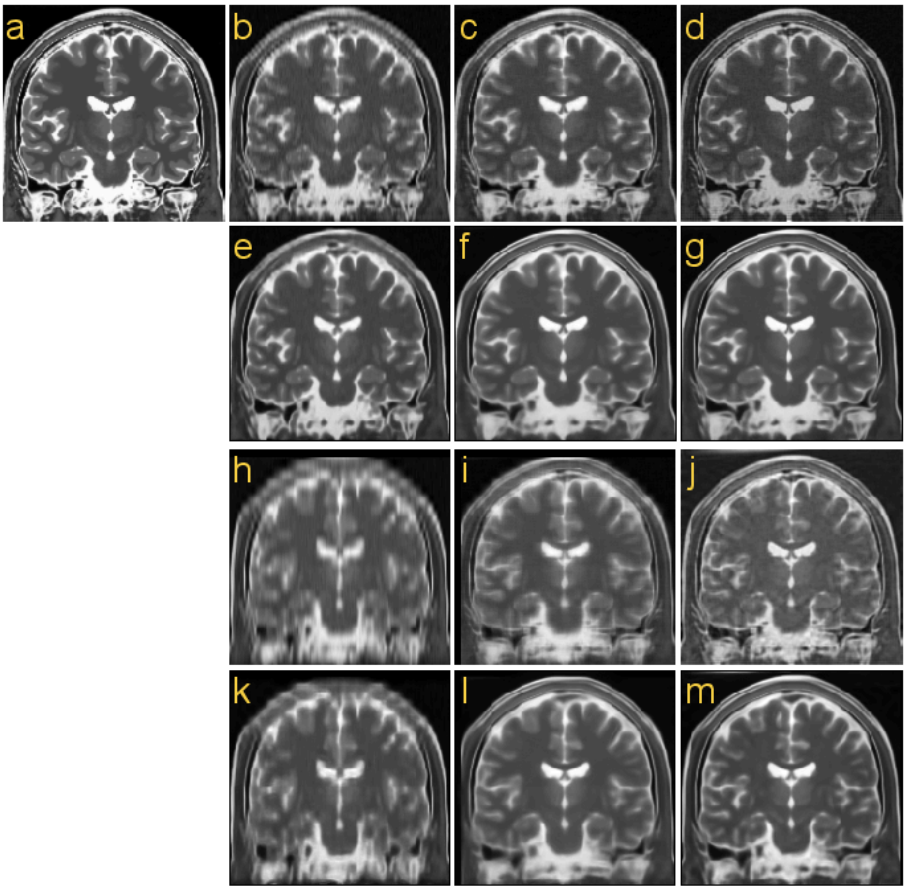



**Fig. 3.** Coronal slice of the simulated brain images reconstructed using six methods: (a) the noise-free HR reference image, (b) cubic interpolation, (c) 3orthoAvg, (d) LASR, (e) siPBSR, (f) 3orthoPBSR, (g) LASR+PBSR. The reconstruction method of (h)-(m) corresponds to (b)-(g). See Section 3.2 for nomenclature. Reconstructions (b)-(g) and (h)-(m) are, respectively, based on stacks of 3 mm and 7 mm slice thickness.

and thinner slices yielded better reconstructions, and in most cases the reconstructions based on three slice stacks were better than the ones based on a single stack. More importantly, compared to interpolation and LASR, PBSR consistently improved the PSNR and mSSIM of the resulting images. In all cases but one, either 3orthoPBSR or LASR+PBSR performed best.

### 3.3 Qualitative Evaluation

Figure 3 shows coronal brain MR images reconstructed from either one or three LR stacks with slice thickness of 3 mm or 7 mm generated from the original HR image with a noise level of 3%. The results of the proposed PBSR method (e-g,

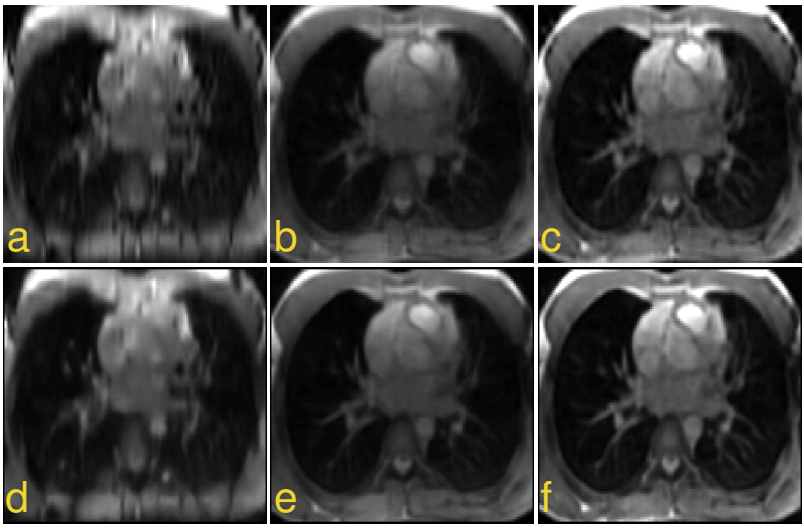



**Fig. 4.** Axial slice of the lung data reconstructed using the six methods: (a) cubic interpolation, (b) 3orthoAvg, (c) LASR, (d) siPBSR, (e) 3orthoPBSR, (f) LASR+PBSR.

k-m) appear both sharper and less noisy than the input images (b-d, h-j). Figure 4 shows axial reconstructions of the lung MRI data. The images demonstrate the potential of multi-image SRR for MRI slice stacks: comparing (a), an interpolated coronal slice stack, with (b, c, e, f), where three stacks are combined using SRR methods, the difference is striking. Comparing the SRR images based on three slice stacks with each other, again, the PBSR images (e, f) appear sharper and less noisy than the images they are based on (b, c). PBSR based on a single slice stack (d) also looks sharper and less noisy than the original interpolated image (a), but in this case it is difficult to assess whether (d) becomes a more accurate representation of the underlying anatomy.

## 4 Discussion and Conclusions

The main contribution of this paper is the idea of exploiting redundancy across scales in an MRI slice stack for SRR. This idea is especially appealing because of the vast amounts of such data that could potentially be enhanced using the proposed method. Cross-scale self-similarity has previously been exploited for SRR in MRI [8, 9], but, to the best of our knowledge, the present study is the first that takes explicit advantage of the relation between structural features in the HR slices of a stack and in the orthogonal LR slices.

Our quantitative evaluation revealed that applying PBSR on an initial HR estimate consistently improves the PSNR and mSSIM. The fact that the PBSR improves the HR estimate of another SRR method (LASR) is worth noting. LASR obtains a regularized maximum likelihood estimated after an iterative pro-

cedure that takes into account the acquisition process while maintaining global data consistency. The improved performance after applying PBSR indicates that the image data contain valuable information that is not exploited by the LASR method but can be exploited using a self-similarity prior.

As future work, we plan to study in detail how patch size and the number of patches in the dictionary affect the performance of our method. The number of blurring levels of the dictionary is another free parameter whose effect on the performance will be thoroughly tested. Such results may, however, not generalize easily, since they will be dependent on the specific anatomy, and on the image quality. Also, we have used only the first NNs for our reconstruction. Using an average of the first $n$ NNs may improve the results. Finally we will study how the size of the used dictionary can be optimally chosen to balance computational cost and reconstruction quality.

## References

1. Plenge, E., Poot, D.H.J., Bernsen, M., Kotek, G., Houston, G., Wielopolski, P., van der Weerd, L., Niessen, W.J., Meijering, E.: Super-resolution methods in MRI: Can they improve the trade-off between resolution, signal-to-noise ratio, and acquisition time? Magn. Reson. Med. 68, 1983–1993 (2012)
2. Elad, M., Datsenko, D.: Example-based regularization deployed to super-resolution reconstruction of a single image. Comput. J. 50, pp. 1–16 (2007)
3. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. IEEE Trans. Pattern Anal. Mach. Intell. 24, 1167–1183 (2002)
4. Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Process. 15, 3736–3745 (2006)
5. Yang, J., Wright, J., Huang, T.S., Ma, Y.: Image super-resolution via sparse representation. IEEE Trans. Image Process. 19, 2861–2873 (2010)
6. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. Proc. CVPR, pp. 60–65 (2005)
7. Glasner, D., Bagon, S., Irani, M.: Super-resolution from a single image. Proc. ICCV, pp. 349–356 (2009)
8. Rousseau, F.: A non-local approach for image super-resolution using intermodality priors. Med. Image Anal. 14, 594–605 (2010)
9. Manjon, J.V., Coupé, P., Buades, A., Fonov, V., Collins, D.L., Robles, M.: Non-local MRI upsampling. Med. Image Anal. 14, 784–792 (2010)
10. Ebrahimi, M., Vrscay, E.: Solving the inverse problem of image zooming using "self-examples". In: Kamel, M., Campilho, A. (eds). Proc. ICIAR, LNSC, vol. 4633, pp. 117–130. Springer, Heidelberg (2007)
11. Protter, M., Elad, M., Takeda, H., Milanfar, P.: Generalizing the non-local means to super-resolution reconstruction. IEEE Trans. Image Process. 18, 36–51 (2009)
12. Cocosco, C.A., Kollokian, V., Kwan, R.K.-S., Evans, A.C.: BrainWeb: Online interface to a 3D MRI simulated brain database. NeuroImage 5, S425 (1997)
13. Poot, D., Van Meir, V., Sijbers, J.: General and efficient super-resolution method for multi-slice MRI. In: Tianzi, J., Nassir, N., Pluim, D., Viergever, M. (eds.) Proc. MICCAI, LNCS, vol. 6361, pp. 615–622. Springer, Heidelberg (2010)
14. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 13, 600–612 (2004)