Article

An Atomic Force Acoustic Microscopy Image Fusion Method Based on Grayscale Inversion and Selection of Best-Fit Intensity

Zhaozheng Chen, Xiaoqing Li and Mingyue Ding *

Molecular Biophysics Key Laboratory of Ministry of Education of China, Department of Biomedical Engineering, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China; u201614392@hust.edu.cn (Z.C.); grace@hust.edu.cn (X.L.)

* Correspondence: myding@hust.edu.cn

Received: 16 October 2020; Accepted: 20 November 2020; Published: 3 December 2020

Abstract: Atomic force acoustic microscopy (AFAM) can provide surface morphology and internal structures of the samples simultaneously, with broad potential in non-destructive imaging of cells. As the output of AFAM, morphology and acoustic images reflect different features of the cells, respectively. However, there are few studies about the fusion of these images. In this paper, a novel method is proposed to fuse these two types of images based on grayscale inversion and selection of best-fit intensity. First, grayscale inversion is used to transform the morphology image into a series of inverted images with different average intensities. Then, the max rule is applied to fuse those inverted images and acoustic images, and a group of pre-fused images is obtained. Finally, a selector is employed to extract and export the expected image with the best-fit intensity among those pre-fused images. The expected image can preserve both the acoustic details of the cells and the background's gradient information well, which benefits the analysis of the cell's subcellular structure. The experiments' results demonstrated that our method could provide the clearest boundaries between the cells and background, and preserve most details from the morphology and acoustic images according to quantitative comparisons, including standard deviation, mutual information, Xydeas and Petrovic metric, feature mutual information, and visual information fidelity fusion.

Keywords: image fusion; grayscale inversion; atomic force acoustic microscope

1. Introduction

High-resolution and nondestructive subcellular imaging instruments ensure that observing the activities of cells accessible is always on demand in biology development. In the 1990s, the introduction of an atomic force microscope (AFM) [1] provided a high-resolution and non-destructive measurement tool at macroscopic levels [2], which is a powerful platform for biological samples, from single molecules to living cells, to be visualized and manipulated [3]. As a near-field microscope, AFM has high resolution in the near field, but it is difficult to detect the inner structures [4].

Atomic force acoustic microscope (AFAM) is a nanoscale microscope that provides ultra-high-resolution images of the sample without destroying the intracellular structures [5]. It is developed by combining AFM with an ultrasound imaging module. The ultrasound transducer produces a single frequency acoustic wave under the cell, and the vibration of the probe is detected by the spot optical tracking system. As the outputs of AFAM, morphology images and acoustic images provide different information: morphology image provides the topography features of the cell by detecting the adhesive between probe and cell, i.e., the AFM image, while the acoustic image shows the phase of the acoustic waves at the same position, which reflects all contributions in the path to make the

acoustic signal shift in the phase. In recent years, AFAM is applied for medical and biological research. For example, the observation of liver cells [6]. AFAM is applied to analyze small cells nowadays. The cells are so thin that signal shift is slight because of the thickness. Hence, the acoustic image only reflects the subcellular structure of the cell. Since the acoustic and morphology image provides completely different types of information, it will be beneficial for studying the cells' structure if we combine them together.

Image fusion is an important research topic in many related areas such as computer vision, remote sensing, robotics, and medical imaging etc. [7]. It is designed to combine multiple source images into a fused image that contains more information than individual source images. Image fusion is widely used in multi-mode image processing, such as visible and infrared images, computed tomography, and magnetic resonance images, and PET-CT images. Strategies for image fusion vary for different fusion tasks. Multi-scale transform (MST) is one of the most popular tools applied in image fusion tasks and has been fully studied. It divides the image into a low-frequency map, which shows a smoothed overall characteristic of the image, and a series of high-frequency maps, which contain the edge details of the image. The commonly used MST methods are the Laplacian pyramid (LP) [8] and nonsubsampling contourlet transform (NSCT) [9]. NSCT is an improved version of contourlet with good fusion results. Unfortunately, its computational efficiency is low. Fusion strategies based on MSTs can be further developed by using an effective way for low-frequency and high-frequency maps fusion to acquire high fusion quality. Yu Liu [10] proposes an image fusion framework based on MSTs and sparse representation (SR) [11], which performs well for the fusion of multi-modality images but needs complex computation. In addition to MST-based methods, deep learning for image fusion has attracted a lot of interest. In recent years, generative adversarial networks (GANs) [12] have been proven to provide more details than the traditional Convolutional Neural Networks (CNNs) in image generation. Jiayi Ma [13] introduced an end-to-end fusion method based on FusionGAN, avoiding manually designing fusion models in traditional methods. However, the network's loss function is still manual designed and the method only works well in visible and infrared images fusion [13]. Although various fusion methods show high performance in medical, multi-focus, and visible-infrared image processing, methods for fusing AFAM images need further research because the observer has difficulty to fuse the different information from the morphological image and acoustic image by mind. Morphology and acoustic images provide different information about the cells, but the high level of the morphology image's intensity may cover the details of the acoustic images. We propose a new method that uses grayscale to efficiently preserve both morphology and acoustic images. It highlights the position and shape of cells and preserves the cell's acoustic details in a single image. First, grayscale inversion is used to transform the morphology images into a series of inverted images with different average intensities. Next, the max rule is applied to fuse the inverted image and acoustic image to obtain a group of pre-fused images. Finally, a selector is employed to determine the expected image with the best-fit intensity among the pre-fused images.

2. Proposed Method

In this section, we introduce our fusion method and details to replicate the experiment.

2.1. Review of the Fusion Method

Figure 1 shows that the proposed fusion method contains three steps: inversion, fusion, and selector. Firstly, a morphology image is inverted into a series of inverted images using different averages. The max rule is then used to fuse the inverted images and acoustic image into the pre-fused images sequences. Finally, a selector is used to choose the expected image from the pre-fused images sequence as the fusion images' output.

Figure 1. Flow chart of the proposed method.

2.1.1. Inversion

Our proposed method's first step is inverting the morphology image into a series of inverted images whose average intensities are different from each other. Our grayscale inversion approach is a simple linear approach, which is defined as below:

$$I_m^k(i, j) = V_{max} - I_m(i, j) + k - 1, \quad (1)$$

where I_m is the original morphology image. I_m^k is the k^{th} inverted morphology image, V_{max} is the largest gray level of I_m . $k = \{1, 2, \dots, L\}$ is the collection of inverted morphology images, and L is the maximum gray value of the acoustic image. We apply greyscale inversion to match the gray level diffusion of morphology and acoustic images. In Figure 2, we show the morphology, inverted, and acoustic image. The morphology image's intensity relates to the sample's height, so the bright area is the cell region, and the dark area is the background, respectively. As for acoustic images, many high-intensity noises surround the cell region, and the acoustic details in the cell are darker than the same region in the morphology image. Grayscale inversion aims to produce a high-intensity background and low-intensity cell region. The high-intensity background will reduce the contract of the noise and even cover it. Simultaneously, the low-intensity cell region makes acoustic details easier to be presented. The bright region, which is considered as background, may be beyond the bounds of the gray level. To present the cell's shape and gradient in the morphology image, the dark areas, which is considered the cell region, should take priority of the fusion work. As a trade-off, we truncate the image gray level to the bounds of the morphological image gray level, i.e., the pixels whose intensity is larger than 255 will be truncated to 255 for the 8-bit image.

Figure 2. (a) Morphology image, (b) inversed morphology image, and (c) acoustic image.

2.1.2. Max Rule

The fusion rule of inverted and acoustic images is the max rule. At the pixel (i,j) , it compares the intensity of the k^{th} inversed image and acoustic image, and selects the larger one as the output. It is defined as follows:

$$I_f^k(i, j) = \max(I_m^k(i, j), I_a(i, j)), \quad (2)$$

where I_f^k is the k^{th} pre-fused image, I_a is the acoustic image. Now, we fuse the information from both morphology and acoustic image. In the background, because the inversed image has a higher average gray level, the pre-fused image has smooth and high-intensity areas. In the boundary of the cell, the gradient from the bright region to dark region indicates the cell's shape. In the cell region, acoustic details are preserved because they have higher intensity than the inversed image. In Figure 3, as the increasing value of k , the pre-fused image provides more gradient information of the inversed morphology image but loses details of the acoustic image. Thus, the next step is to choose the image that preserves both morphology and acoustic information. We name this ideal image as the best-fit intensity image.

Figure 3. The pre-fused images with different k values: (a) $k = 0$, (b) $k = 50$, (c) $k = 100$, (d) $k = 150$, (e) $k = 200$, (f) $k = 250$.

2.1.3. Selector

To choose the best-fit intensity image, a selector is designed to measure the information loss of the pre-fused images to select the best-fit intensity image as the final output of the fused image.

Because AFAM fusion aims is to combine the information of the cell, the regions of the cells should be extracted, and the fusion problem could be simplified to the fusion problem in the regions of cells. To acquire the position and shape of the cell, a cell extraction algorithm based on Otsu's approach [14] was developed. Erosion and dilation steps are used to reduce the influence of the noise and smoothen the boundary of the cell region. Our cell extraction method is a triple-step process. The first step is detection. Otsu's threshold is used to segment the morphology image into CEL_0 and BKG at gray level t , which is defined as follows:

$$BKG = \{0, 1, 2, \dots, t\} , \quad (3)$$

$$CEL_0 = \{t + 1, t + 2, \dots, L - 1\} , \quad (4)$$

where t is the threshold of classifying background and the cell. To clarify, Otsu's threshold is not an accurate segmentation method because it does not perform well in segmenting smooth boundary. Our method is not sensible to the segmentation accuracy.

However, the small object, which is not the region of interest in the view of content but has the same gray level section, does harm to the fusion performance. To clear them, we apply the erosion step. We assume that the main part of the morphology image is the cell, and noise occupies small and separated regions. To clear these areas, the erosion step is defined as follows:

$$CEL_1 = I_m \ominus (BKG \oplus A) , \quad (5)$$

$$L_A = \text{round}\left(\sqrt{\sum(CEL_0) \times (\sqrt{p_1} - 1)}\right) , \quad (6)$$

where A is a disk-shaped structuring element, L_A is the size of A , and $p_1 = 1.2$ is the multiple coefficient that determines the narrowing of the original segmentation of the cell. The function of A is to clear the noise regions which have a similar intensity feature of the cells, and the proper value of p_1 is to clear all the noise regions but still preserve the region of the cells. According to our hypothesis, the noise regions are removed during the erosion process but the cell region is narrowed either. Therefore, the dilation step, which is the third step, is used to recover the cell region. It is defined as follows:

$$CEL = CEL_1 \oplus B , \quad (7)$$

$$L_B = \text{round}\left(\sqrt{\sum(CEL_1) \times (1 - \sqrt{p_2})}\right) , \quad (8)$$

where B is a disk-shaped structuring element, L_B is the size of B , and p_2 is a multiple coefficient that determines the extension of CEL_1 . The function of B is to recover the region of cells and make the boundary of the cells smooth. The value of p_2 is determined by p_1 , and it is defined as follows:

$$p_2 = \frac{1}{p_1} . \quad (9)$$

At last, we obtain CEL , which represents the region of cell, and BKG , representing the background. The relationship of morphology image, BKG and CEL is defined as follows:

$$I_m = \{CEL, BKG\} . \quad (10)$$

Next, to set the standard of the best-fit intensity image among the pre-fused images, we define a multi-regional information loss (MRIL) to measure the information loss. It is a double-stage algorithm, which would calculate the total information loss of the inner and outer cells. To define the set of the

criterion of the pre-fused images, characteristic of inverted morphology and acoustic image should be taken into consideration. As it is shown in Figure 2, inverted morphology and acoustic images show different information in different regions. In the cell region, because of grayscale inversion, the intensities of inverted morphology are weaker than those in the background and provide little information about the inner structures. By contrast, the acoustic image provides much richer details in the region of the cell. Thus, the fused image should preserve as many details of the acoustic image as we can. In terms of background, inverted morphology image is used to show the cell's shape and position, which is not clear in acoustic image. To measure the total information loss, MRIL is defined as follows:

$$L_{total}^k = \frac{L_{CEL}^k + L_{BKG}^k}{2}, \quad (11)$$

where L_{CEL}^k is the loss of cell and L_{BKG}^k is the loss of background. They are given by:

$$L_{CEL}^k = \frac{\sum_{i,j \in CEL} |I_f^k(i,j) - I_{ac}(i,j)|}{\Sigma(CEL)}, \quad (12)$$

$$L_{BKG} = rSFe(I_f^k(i,j), I_m^k), \quad (13)$$

where I_f^k is the k^{th} pre-fused image, I_{ac} is the acoustic image, $rSFe(I_f^k, I_m^k)$ has a similar definition of the ratio of SF error [15]. The difference is that our rSFe method measures the space-frequency error between the pre-fused image and inverted morphology image in the region of BKG instead of fused image and source images. The reason why we choose rSFe to measure the loss in the background for the following two reasons. Firstly, the space-frequency remains the same no matter what the value of k is. Secondly, rSFe is a suitable regulation that makes the background as bright as possible to cover the acoustic noise, and it is not sensible to the segmentation accuracy of the cell either. Given that the acoustic image is characterized by their pixel intensities, MRIL is sensitive to among the pixels preserved from the acoustic image in the region of the cell.

According to the definition of information loss, the expected image is the image with the lowest L_{total} . Thus, the selected image is defined as follows:

$$I_f = I_f^{k_0}. \quad (14)$$

2.2. Experiment

In this section, we introduce the experiment of our method, including data acquisition, performance metrics, and the results and discussion of the experiment.

2.2.1. Data Acquisition

All the AFAM images we used in the experiment are acquired by CSPM5500, which is an open atomic force acoustic microscope from Guangzhou Benyuan Nano-Instruments Ltd., Guangzhou, China. *Escherichia coli* (*E. coli*) and *Staphylococcus aureus* (*S. aureus*) are the AFAM measurements samples. They are stored by the China Center for Type Culture Collection in Wuhan (Wuhan H22). As shown in Figure 4, 16 pairs of source images are employed to verify the effectiveness of the competing methods. Among them, there are 15 pairs of *S. aureus* images, and there is 1 pair of *E. coli* images. Four fusion methods, including Laplacian pyramid (LP) [10], nonsubsampling contourlet transform (NSCT) [11], fusion method based on Laplacian pyramid and sparse representation (LP-SR) [10], and FusionGAN [13], are applied for the comparison of the fusion performance. According to the original papers, we set the parameters of competing methods but use our AFAM images as the source images. Meanwhile, since visible and infrared images fusion has a similar pattern to AFAM images, we use our AFAM images database to retrain the FusionGAN to obtain better fusion performance at AFAM fusion task.

Figure 4. Sixteen pairs of morphology and acoustic images.

2.2.2. Performance Metrics

Because the reference image does not exist, it is hard to quantitatively evaluate a fused image's quality. In recent years, many fusion metrics have been proposed to evaluate fusion methods' performance, but none of them are widely believed. Therefore, it is necessary to apply several fusion metrics to evaluate fusion methods. To better evaluate the fusion performance of the cell region, we manually segment the source image of the source image into cell, and background and all the evaluation metrics measure the fusion performance in the areas of cells. Segmentation results are shown in Figure 5. The quantitatively evaluate metrics, we chose the following five metrics, including standard deviation (SD) [16], mutual information (MI) [17], Xydeas and Petrovic metric ($Q^{AB/F}$) [18], feature mutual information (FMI) [19], and the visual information fidelity fusion (VIFF) [20]. Here are the definitions of the evaluation metrics.

1. SD is a value of the degree of a set of image data averages, reflecting the fused image's contract. Mathematically, SD is defined as follows:

$$SD = \sqrt{\sum_{i=1}^M \sum_{j=1}^N (I_F(i, j) - \mu)^2} , \quad (15)$$

where I_F is the fused image, and μ is the mean value of the fused image.

2. MI measures the degree of dependency between two events A and B . It is defined as follows:

$$I_{UV}(U, V) = \sum_{v \in V} \sum_{u \in U} p(u, v) \log_2 \frac{p(u, v)}{p(u)p(v)} , \quad (16)$$

where $p(u, v)$ is the joint distribution and $p(u)$ and $p(v)$ are the marginal distribution. Considering two source images A and B , and the fused image F , MI is defined as follow:

$$MI = I_{FA}(f, a) + I_{FB}(f, b) . \quad (17)$$

3. $Q^{AB/F}$ is an objective performance metric which measures the relative amount of edge information that is transferred from the source images to the fused image. It is defined as follows:

$$Q_{\frac{AB}{F}} = \frac{\sum_{i=1}^N \sum_{j=1}^M (Q^{AF}(i, j)w^A(i, j) + Q^{BF}(i, j)w^B(i, j))}{\sum_i \sum_j (w^A(i, j) + w^B(i, j))} , \quad (18)$$

$$Q^{AF}(i, j) = Q_g^{AF}(i, j)Q_o^{AF}(i, j) , \quad (19)$$

$$Q^{BF}(i, j) = Q_g^{BF}(i, j)Q_o^{BF}(i, j) , \quad (20)$$

where $Q_g^{XF}(i, j)$ and $Q_o^{XF}(i, j)$ are the edge strength and orientation preservation value at the pixel (i, j) . $w^A(i, j)$ and $w^B(i, j)$ reflect the weight of $Q^{AF}(i, j)$ and $Q^{BF}(i, j)$.

4. FMI computes the amount of feature transfer from source images to the fused image. A gradient map of an image provides information of texture, edge strength, and contrast included.
5. VIFF is a recently proposed metric that measures the visual information fidelity between the source images and fused image. It measures the visual information of every block in each sub-band of the fused image based on the Gaussian scale mixture model, distortion model, and human visual system (HVS) model.

Figure 5. Segmentation result of the compared images. The red circles are the regions of cells.

3. Results and Analysis

Firstly, we try to compare the fusion quality of different methods. As for Figure 6, three typical images are selected to intuitive results on the fusion performance. The first and second rows show the fusion method's ability to provide the details between the acoustic image. The third row requires the fusion method to make a distinction of two cells. The first two columns are morphology and acoustic images, the last column presents the fusion result of our method. The other columns are the results of the compared fusion methods. From the result, it is easy to conclude that our fused images are the best at preserving acoustic information. The inverted morphology part smartly highlights the shape of cell and avoids covering the acoustic details. LP and RP present similar fusion results. They can both acquire boundary and the acoustic information of cells, but they lack the acoustic information inside the cell. The fusion result of NSCT shows that the edge of the inner structures is sharp, and the shape of cells is easy to distinguish, but the contrast of the inner structure is low. FusionGAN is good at preserving morphology information, but it only keeps gradient information of acoustic images. As a result, the internal structures of the cells remain low contrast and hard to be found. Our method's fusion result is much different from the others, which is the result of grayscale inversion applied to morphology images. The area outside the cells is not the region of interest, so the strong intensity of the inverted morphology image would cover the acoustic noise, which is why the region of the background is bright. The cells' region is thought to contain much acoustic information, and the low insensitivity of morphology information has high contract with the acoustic information and highlights the detail of the internal structures of cells.

Figure 6. Fusion results on three typical morphology and acoustic image. (From **left to right**: morphology image, acoustic image, Laplacian pyramid (LP), nonsubsampling contourlet transform (NSCT), fusion method based on Laplacian pyramid and sparse representation (LP-SR), fusion GAN, proposed).

Moreover, the gradient from the bright background and the dark region of the cell forms the boundary. As a result, the obtained results by the proposed have sharp edges, more details, and enhanced contrast inside the internal structure and clear boundary of cells. The analysis of fusion images' features indicates that the application of grayscale inversion significantly highlights the cell and prevents the bright morphology pixel to cover the cell structure.

To assess the fusion quality in the cells' region, we further give quantitative comparisons of the five methods. We first manually segment the morphology image to obtain the weight map of the cell. Then, we apply the weight map to pre-process the fused image, morphology image, and acoustic image. Next, the five fusion metrics are used to evaluate the fusion quality of the pre-processed fusion image. Tables 1–5 list the metrics of the 16 images; the 9th image is *E. coli* and the others are *S. aureus*. Table 6 shows the result of the *t*-test between our proposed method and the methods with the second-highest value. According to Tables 1–3, our method approximately has the most extensive MI, $Q^{AB/F}$, and FMI. The largest MI implies that our method preserves the most information of the source images. Our method has the best MI because it is enhanced to save the intensities of the acoustic image in the region of the cell, which is much stronger than the morphology information. The largest $Q^{AB/F}$ demonstrates that our fused image has rich edges information. Since the acoustic image provides most of the edge information, we can also consider that our fused image has the most considerable acoustic information. The largest FMI also demonstrates that our method has the most features including texture, edge strength, and contrast. Although the values of VIFF and SD do not have as large an advantage as the above three metrics, they still have the highest average among the compared methods. The results of the *t*-test imply the significance of our method. The result of VIFF yet demonstrates that our method is great, consistent with the human visual system. Although the VIFF of NSCT is close to our method, other metrics imply a certain gap between NSCT and the proposed method. The result of SD shows that our fused image expressed high contrast but not clear enough to be considered as the best. According to Table 5, FusionGAN shows much better fusion performance than the other method in the 10th image, and results in low reliability according to the *t*-test shown in Table 6. Further research implies that SD of our method is highly dependent on acoustic image. The low SD of the fused image is the result of low contrast of the 10th acoustic image. The fusion metrics above indicate that our method is good at preserving acoustic detail inside the cell, which is beneficial to analyzing the nanoscale and subcellular structure of the cell.

Table 1. Mutual information (MI) between proposed and compared methods. The bold represents the best value.

Image	LP	NSCT	LP-SR	FusionGAN	Proposed
1	3.4138	3.5749	3.7945	4.4764	7.7349
2	3.8354	4.2508	4.1686	4.8191	8.0714
3	3.7562	4.0227	3.7419	4.7124	7.9157
4	4.3860	4.9246	4.6630	5.3442	8.8741
5	3.4103	3.8163	3.8791	4.3792	8.7135
6	4.0918	4.5772	4.3210	4.5940	7.8550
7	3.7621	4.2687	4.0084	4.7183	7.8393
8	3.3896	4.3825	4.3964	4.7479	8.1814
9	2.2528	2.3980	2.6728	3.6253	7.0375
10	3.1164	4.0883	3.1396	5.3285	6.3951
11	3.9045	4.0335	4.2035	4.4849	7.7655
12	3.4848	3.8969	3.7320	4.1230	8.0500
13	4.2784	4.5256	4.1659	4.5776	8.0357
14	4.8693	5.3385	4.7021	5.1016	8.8752
15	3.5955	3.8034	3.7454	3.9664	7.9100
16	3.4814	3.7239	3.9381	3.9120	8.3066

Table 2. Xydeas and Petrovic metric ($Q^{AB/F}$) between proposed and compared methods. The bold represents the best value.

Image	LP	NSCT	LP-SR	FusionGAN	Proposed
1	0.3402	0.4143	0.3211	0.1969	0.6621
2	0.3506	0.4801	0.3490	0.2113	0.7311
3	0.3383	0.4259	0.3068	0.2457	0.6585
4	0.3116	0.4388	0.2884	0.2585	0.6331
5	0.3978	0.5288	0.3846	0.2410	0.6962
6	0.3695	0.5073	0.3417	0.2488	0.7218
7	0.3437	0.4264	0.3379	0.2317	0.7033
8	0.3885	0.4846	0.3971	0.2116	0.7361
9	0.3552	0.3298	0.3597	0.1052	0.6111
10	0.3217	0.5089	0.3352	0.2033	0.6087
11	0.2937	0.4144	0.2804	0.1764	0.5886
12	0.3587	0.5032	0.3679	0.3092	0.6906
13	0.3007	0.4015	0.3057	0.2616	0.5988
14	0.3207	0.4383	0.3246	0.2190	0.5420
15	0.3995	0.5233	0.4140	0.2292	0.6239
16	0.4181	0.5515	0.4269	0.2641	0.7636

Table 3. Feature mutual information (FMI) between proposed and compared methods. The bold represents the best value.

Image	LP	NSCT	LP-SR	FusionGAN	Proposed
1	0.3058	0.3343	0.2972	0.2452	0.5037
2	0.2888	0.3258	0.2855	0.2525	0.5203
3	0.2962	0.3229	0.2851	0.2548	0.5122
4	0.3555	0.3821	0.3381	0.3038	0.5323
5	0.3501	0.3752	0.3451	0.2977	0.5412
6	0.3114	0.3447	0.3030	0.2646	0.5212
7	0.3227	0.3558	0.3145	0.2537	0.5402
8	0.2877	0.3254	0.2847	0.1620	0.5330
9	0.3176	0.3305	0.3172	0.2837	0.5121
10	0.2619	0.3228	0.2676	0.2571	0.5686
11	0.3115	0.3541	0.3107	0.1512	0.4802
12	0.3166	0.3559	0.3151	0.3400	0.4911
13	0.3195	0.3540	0.3175	0.3290	0.4784
14	0.3652	0.3996	0.3647	0.3530	0.5091
15	0.3175	0.3529	0.3175	0.3231	0.4554
16	0.3119	0.3421	0.3119	0.3115	0.4884

Table 4. The visual information fidelity fusion (VIFF) between proposed and compared methods. The bold represents the best value.

Image	LP	NSCT	LP-SR	FusionGAN	Proposed
1	0.6500	0.7718	0.6699	0.5156	0.9963
2	0.6620	0.9232	0.6685	0.6199	1.1875
3	0.6651	0.7690	0.6993	0.6375	1.1042
4	0.6088	0.7561	0.6281	0.5929	1.0091
5	0.8627	0.9899	0.9435	0.6675	1.1277
6	0.7850	0.9653	0.8214	0.6810	1.1944
7	0.6903	0.8971	0.6923	0.5714	1.2731
8	0.8110	1.0376	0.9095	0.6186	1.2405
9	0.6954	0.6459	0.6968	0.5432	0.8411
10	0.7083	1.1616	0.6935	1.2347	1.2102
11	0.6955	0.7798	0.7026	0.4981	0.9608
12	0.7036	0.8571	0.7122	0.7240	1.0765
13	0.7189	0.7677	0.7398	0.6632	0.9949
14	0.675	0.7842	0.6979	0.5911	0.8309
15	0.7891	0.8699	0.7959	0.6262	1.0216
16	0.7700	0.9220	0.7949	0.7075	1.1852

Table 5. Standard deviation (SD) between proposed and compared methods. The bold represents the best value.

Image	LP	NSCT	LP-SR	FusionGAN	Proposed
1	36.9658	39.8320	39.9261	34.5990	40.1020
2	38.4776	47.6098	47.4358	43.2366	50.6317
3	37.7895	37.6134	41.4245	33.5233	42.3467
4	41.5586	46.5872	47.7077	41.2347	50.8664
5	31.9550	42.2352	37.7491	39.0611	48.8546
6	33.1649	39.8875	41.0251	33.8348	45.9419
7	28.9697	35.6636	33.4225	34.4395	41.0504
8	33.3158	41.9951	42.7390	46.4726	47.0966
9	37.9474	35.4992	31.9065	29.4996	34.1644
10	9.4739	18.4304	8.5953	41.7376	12.6174
11	40.2326	44.4862	42.3177	42.0241	38.2377
12	33.8317	39.4966	36.3585	41.0672	44.4249
13	38.0081	40.5067	38.7161	37.3651	42.8768
14	43.4355	49.6466	43.2625	41.2975	44.2562
15	36.2075	37.5342	39.7434	38.7902	45.0851
16	35.6061	39.3488	42.3115	43.2114	48.3688

Table 6. T test between proposed and compared methods.

	MI	$Q^{AB/F}$	FMI	VIFF	SD
LP	1.38×10^{-15}	4.17×10^{-14}	8.07×10^{-12}	9.99×10^{-9}	1.83×10^{-4}
NSCT	9.14×10^{-14}	8.71×10^{-10}	3.35×10^{-11}	8.38×10^{-8}	5.46×10^{-2}
LP-SR	7.01×10^{-17}	2.64×10^{-13}	5.51×10^{-12}	4.07×10^{-8}	5.42×10^{-4}
FusionGAN	1.02×10^{-11}	3.25×10^{-14}	7.40×10^{-10}	4.45×10^{-8}	1.64×10^{-1}

To compare the computationally intensive, we list the operation time of processing a pair of images in Table 7. Since FusionGAN runs in TensorFlow, it is not suitable to compare with other methods running in MATLAB. Our method is slower than LP-SR and LP, but it can still fuse the images in less than one second.

Table 7. Operation time of the fusion methods.

Method	Time (s)
NSCT	1.114156
LP	0.005338
LP-SR	0.040919
Proposed	0.4031

4. Conclusions

In this paper, a novel fusion method for AFAM image fusion was proposed based on grayscale inversion and best-fit intensity selection. Grayscale inversion aims to change the grayscale feature of the morphology image and avoid the tradeoff between morphology and acoustic information. The inversion process is improved by inverting a single image into a series of inverted images and

selecting the most suitable one as the output, which could avoid the tradeoff between high-intensity morphology information and complex acoustic details. The image segmentation method based on Otsu's approach segments the regions of cells and background, which can greatly help reduce the interruption of the morphology image and improve segmentation precision. The selection of the best-fit intensity is applied to measure the total information loss of the pre-fused image and choose the best one as the output of our method, and MRIL is used as the fusion parameter of optimization. Experiments on different fusion methods can prove that our method is the best at presenting the features of source images, and saving information from the AFAM images. The results of quantitative comparisons also demonstrated that the contrast of our fused images depends on the source of the acoustic images, which means that our method is designed to restore acoustic image in the region of the cell. The results indicate that our method is good at highlighting the cell and providing subcellular structure, which is beneficial for the analysis of the cell's structure. Currently, the quality of acoustic image limits the fusion performance, while in the future, we will focus on the fusion rules for the inverted and acoustic images and make our fusion model more robust and efficient.

Author Contributions: Conceptualization, X.L. and Z.C.; methodology, Z.C.; investigation, Z.C. and X.L.; validation, Z.C.; data curation, Z.C.; resource, X.L.; writing original draft preparation Z.C.; writing-review and editing, X.L. and M.D.; supervision, M.D.; project administration M.D.; funding acquisition, M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (grant No. 81571754), and partly supported by the Major National Key Scientific Instrument and Equipment Development Project (grant No. 2013YQ160551) and National undergraduate Innovation and Entrepreneurship Research Practice Program, No. 201918YB05.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shea, S.J.O.; Welland, M.E. Atomic force microscope. *Phys. Rev. Lett.* **1992**, *59*, Unit 2C.2.
2. Ikai, A.; Afrin, R.; Saito, M.; Watanabe-Nakayama, T. Atomic force microscope as a nano- and micrometer scale biological manipulator: A short review. *Semin. Cell Dev. Biol.* **2018**, *73*, 132–144. [[CrossRef](#)] [[PubMed](#)]
3. Dufrêne, Y.F.; Ando, T.; Garcia, R.; Alsteens, D.; Martinez-Martin, D.; Engel, A.; Gerber, C.; Müller, D.J. Imaging modes of atomic force microscopy for application in molecular and cell biology. *Nat. Nanotechnol.* **2017**, *12*, 295–307. [[CrossRef](#)] [[PubMed](#)]
4. Rabe, U. Atomic force acoustic microscopy. *Nanosci. Technol.* **2006**, *15*, 1506–1511.
5. Wang, T.; Ma, C.; Hu, W. Visualizing subsurface defects in graphite by acoustic atomic force microscopy. *Microsc. Res. Tech.* **2017**, *80*, 66–74. [[CrossRef](#)] [[PubMed](#)]
6. Li, X.; Lu, A.; Deng, W.; Su, L.; Wang, J.; Ding, M. Noninvasive subcellular imaging using atomic force acoustic microscopy (AFAM). *Cells* **2019**, *8*, 314. [[CrossRef](#)] [[PubMed](#)]
7. Pure, A.A.; Gupta, N.; Shrivastava, M. An overview of different image fusion methods for medical applications. *Int. J. Sci. Eng. Res.* **2013**, *4*, 129–133.
8. Burt, P.J.; Adelson, E.H. The laplacian pyramid as a compact image code. *IEEE Trans. Commun.* **1983**, *31*, 532–540. [[CrossRef](#)]
9. Chen, X.W.; Ge, J.; Liu, J.G. Non-subsampled contourlet texture retrieval using four estimators. *Appl. Mech. Mater.* **2012**, *263*, 167–170. [[CrossRef](#)]
10. Liu, Y.; Liu, S.; Wang, Z. A general framework for image fusion based on multi-scale transform and sparse representation. *Inf. Fusion* **2015**, *24*, 147–164. [[CrossRef](#)]
11. Yang, B.; Li, S. Multifocus image fusion and restoration with sparse representation. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 884–892. [[CrossRef](#)]
12. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.C.; Bengio, Y. Generative Adversarial Networks. *arXiv* **2014**, arXiv:abs/1406.2661. [[CrossRef](#)]
13. Ma, J.; Yu, W.; Liang, P.; Li, C.; Jiang, J. FusionGAN: A generative adversarial network for infrared and visible image fusion. *Inf. Fusion* **2019**, *48*, 11–26. [[CrossRef](#)]
14. Jong, G.; Hendrick, H.; Wang, Z.; Kurniadi, D.; Aripriharta, A.; Horng, G. Implementation of Otsu's method in vein locator devices. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2018**, *8*, 743–748. [[CrossRef](#)]

15. Zheng, Y.; Essock, E.A.; Hansen, B.C.; Haun, A.M. A new metric based on extended spatial frequency and its application to DWT based fusion algorithms. *Inf. Fusion* **2007**, *8*, 177–192. [[CrossRef](#)]
16. Rao, Y.J. In-fibre Bragg grating sensors. *Meas. Sci. Technol.* **1997**, *8*, 355. [[CrossRef](#)]
17. Ma, J.; Ma, Y.; Li, C. Infrared and visible image fusion methods and applications: A survey. *Inf. Fusion* **2018**, *45*, 153–178. [[CrossRef](#)]
18. Xydeas, C.S.; Petrovic, V.S. Objective image fusion performance measure. *Electron. Lett.* **2000**, *36*, 308–309. [[CrossRef](#)]
19. Asha, C.S.; Lal, S.; Gurupur, V.P.; Saxena, P.P. Multi-modal medical image fusion with adaptive weighted combination of NSST bands using chaotic grey wolf optimization. *IEEE Access* **2019**, *7*, 40782–40796. [[CrossRef](#)]
20. Li, S.; Kang, X.; Fang, L.; Hu, J.; Yin, H. Pixel-level image fusion. *Inf. Fusion* **2017**, *33*, 100–112. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).