Slice Interpolation for Medical Image based on Spatial Geometry Polynomial Fitting

Hui Liu[1] (iD) , Yuxiu Lin[2] (iD) , Shanshan Li[3] (iD) , and Caiming Zhang[4] (iD)

[1]Shandong University of Finance and Economics, liuh_lh@sdufe.edu.cn
[2]Shandong University of Finance and Economics, 15628791525@163.com
[3]Shandong University of Finance and Economics, 2608571951@qq.com
[4]Shandong University, czhang@sdu.edu.cn

Corresponding author: Hui Liu, liuh_lh@sdufe.edu.cn

Abstract. Nowadays, medical image plays an important role to effectively assist clinicians in disease diagnosis and treatment process. However, the image resolution in the axial spatial direction is obviously lower than the cross section direction due to the influence of radiation doses and scanning times. This phenomenon will adversely affect the accuracy of computer-aided diagnosis (CAD). Thus, it is necessary to focus on the research of medical image slice interpolation. In this paper, a fast and flexible volumetric interpolation of method based on spatial geometry polynomial fitting is presented, in order to obtain the high resolution reconstruction results in any direction and proportion. Based on the similarity between medical sequence images, three consecutive CT slices in 3D space are used to reversely reconstruct the fitted space geometry that approximates the original scene, while resampling the space geometry to generate new voxels. First, the reverse sampling on original slice sequence is adopted to locally construct a ternary quadratic polynomial space geometry, which is used to generate the unit space geometry by weighted average. Then, all unit space geometries are pieced together to fit the space geometry. Finally, the spatial geometry is resampled to reconstruct the inter-layer slice. In our experimental results, the proposed method performs the best in average gradient, RMSE, DSSIM and time complexity metrics, allowing for better visual evaluation.

1 INTRODUCTION

As an active research topic, image interpolation and super-resolution reconstruction have achieved widespread use in various branches of image processing, especially in the field of medical imaging [16, 20]. Medical imaging technologies, such as computational tomography (CT), Magnetic Resonance Imaging (MRI) and Positron Emission Computed Tomography (PET), are essential to disease diagnosis and surgery planning. For example, CT imaging method usually expresses the acquired tomographic medical image data as a set of slice sequences. However, the acquisition of CT images is accompanied by high dose radiation. To decrease the radiation amount received by the patients, it is a common practice to reduce the sampling rate and improve the scanning speed, which resulting in the loss of some valuable temporal information and remarkably large slice interval. As a result, most medical imaging volumes are taken anisotropically with a high intra-slice resolution and a low inter-slice resolution. This asymmetry in the resolution leads to problems such as clearly step-shaped in 3D reconstructed models and rough or even broken tissue boundaries, which will undoubtedly affect the accuracy of 3D visualization or image analysis. As such, an accurate and reliable method to upsample the low inter-slice resolution, we refer to as the slice interpolation techniques [2, 11], is much needed.

In addition to generate the accurate 3D reconstructions, medical slice interpolation can also be widely used in medical image segmentation [22], multi-frame super-resolution (MFSR)[15, 12] reconstruction where the aim is to reconstruct high resolution images from their corresponding low-resolution sequences. By adding new virtual slices between two consecutive images, as depicted in Figure 1, the number and information of experimental data sets are increased, in order to boost MFSR and medical image segmentation accuracy. Especially, increasing the amount of training samples is indispensable for the hottest research method, neural network method. Therefore, it is necessary to develop image interpolation techniques, more specifically medical image slice interpolation techniques, to increase the axial spatial resolution of the data acquired by medical imaging modalities.
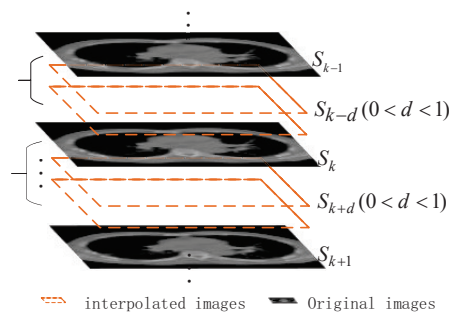

Figure 1: Inter-layer interpolation in fault medical images.

In this paper, we propose a method that generates inter-layer slices based on spatial geometry polynomial fitting, which can reconstruct an arbitrary number of intermediate medical slices from two consecutive slices. An overview of our method is illustrated in Figure 2. First, the reverse sampling on original slice sequence is adopted to locally construct a ternary quadratic polynomial space geometry, which is used to generate the unit space geometry by weighted average. Then, all unit space geometries are pieced together to fit the space geometry. Finally, the spatial geometry is resampled to reconstruct the inter-layer slices. This method was used to increase the temporal resolution in CT sequence image and compared to the state-of-the-art interpolation methods. We further conducted an ablation study to demonstrate the effectiveness of our motion network.

In summary, the contributions are summarized as following:

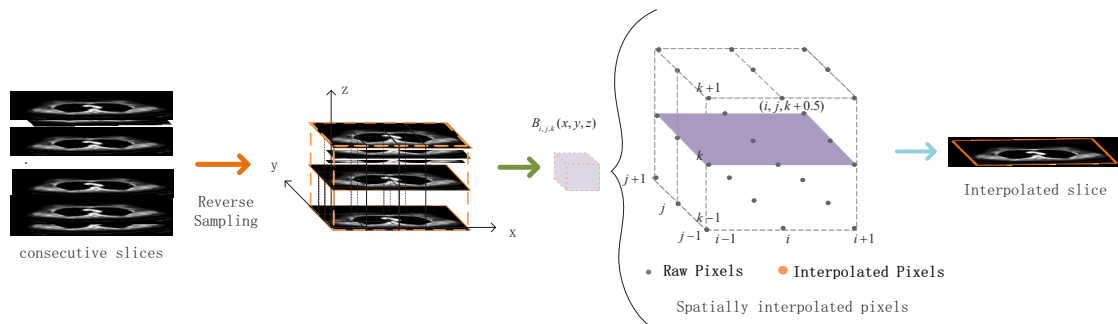a) A fast and flexible medical image volumetric interpolation method is proposed based on spatial geometry

Figure 2: Algorithm framework.

polynomial fitting theory, which can fast upsample the inter-layer slices from anisotropic volumes. The method can be directly applied to interpolate images in arbitrary orientation and scale.

b) The theory of intra-layer pixel interpolation in cross image scale-down is applied to voxel data reconstruction in 3D space. Unlike the traditional model that uses two consecutive images as input, it uses three adjacent images as input to predict the motion trajectory information of the inter-layer image;

c) Using the relevant theory of computer graphics, the spatial geometry of three adjacent layer images is fitted in 3D space, and then resampled to generate new voxels to reconstruct the inter-layer slices. The time complexity is greatly reduced while achieving high quality results.

2 RELATED WORKS

Medical image slice interpolation technology has an important research significance and has spawned a lot of methods [21]. The methods for this task can be categorized into four groups: grayscale-based, shape-based, alignment-based and learning-based methods.

Grayscale-based interpolation methods [10, 9, 23] directly use the grayscale information of the upper and lower images in the original sequence, to interpolate the inter-layer images by a set of basis functions. Nearest neighbor interpolation, linear interpolation, and cubic B spline function interpolation [7] are the common types of such interpolation methods. In the case of small layer spacing, grayscale based interpolation methods are widely used because they are computationally simple, low time complexity and easy to implement. However, for CT image sequences with large structural deformation between adjacent slices, the interpolated images obtained by this method are usually too smooth and contain much artifacts, which cannot accurately reflect the natural gradient of tissue contours. Subsequently, Goshtasby [4] proposed a matching interpolation method, which improved the edge blurring problem to a certain extent and improved the accuracy of the interpolated images, but with high time complexity.

Shape-based interpolation methods [18, 5] generate the contours of the slices interpolated directly from the shapes of original slices in the upper and lower layers. Compared with the grayscale based interpolation methods, it can effectively eliminate artifacts and improve the quality of interpolated images. However, the shape-based method still has significant limitations. On the one hand, it requires a high quality of tissue contour, which is suitable for the case where the upper and lower slices are very similar. If the layer spacing is too large or there are large structural differences between two consecutive CT images, it will seriously affect the effectiveness of this method and lead to large errors in the interpolation results. On the other hand, the extraction and representation of contours make this method complicated and reduce the efficiency of the algorithm.

The alignment-based interpolation methods [17, 13, 2] are widely used to match the local anatomical structures by finding the correspondence between adjacent slices, and to obtain the deformation information of the pixel points to be interpolated based on the found deformation fields to achieve inter-layer interpolation. The alignment-based interpolation can eliminate boundary artifacts while taking into account the various non-rigid deformations of human organ structures, which can produce visually satisfying results. The popular non-rigid alignment methods can be divided into two categories: alignment based on spatial transformations, such as the B-sample-based free form alignment method proposed by Rueckert et al [19], and alignment based on physical models, such as the alignment method based on optical flow estimation [8, 14]. However, the alignment-based interpolation methods have approximate calculations in the alignment process, which may result in multiple pixels in the original image mapped to the same location in the interpolated image, or anomalies in the new interpolated image where there are no pixels mapped to the original image, i.e., missing pixels. Compared with the shape-based and grayscale based interpolation methods, their computational complexity is higher.

In recent years, the rapid development of deep learning has enabled neural network to bring new break-throughs on medical image slice interpolation [6, 16]. Learning-based, especially deep learning-based slice interpolation techniques have a good development prospect. However, its inherent unpredictability and un-interpretability in handling sophisticated tasks are difficult to meet in most practical application scenarios. For example, medical images remain to be labled, and in addition, deep learning-based methods rely on large amounts of training data.

In order to solve the problems of low interpolation accuracy or low efficiency in above, this paper proposes a method of slice interpolation for medical image based on spatial geometry polynomial fitting, which can improve the computational efficiency while obtaining better reconstructed inter-layer images quantitatively and qualitatively.

3 METHOD

3.1 Basic Constraint Condition

The inter-slice interpolation method based on spatial geometric polynomial fitting depends on two constraints:

Condition 1: the pixels between the inter-slice image to be interpolated and the original image sequence all change continuously, and they contain similar structures.

Condition 2: the similarity between the inter-layer image to be reconstructed and the original two consective images is inversely proportional to the distance between the two images.

Suppose the distance between two consecutive images S_k, $S_{k+1} (0 < k < t)$ is 1, and the image $S_{k+d} (0 < d < 1)$ to be interpolated consists of $m \times n$ pixels $P_{i,j,k+d}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n$, which can be regarded as a sampling value on a three-dimensional geometry $F(x, y, z)$. To facilitate discussion, each pixel $P_{i,j,k+d}$ is assumed to be sampled from a unit volume, that is,

$$P_{i,j,k+d} = \int_{k+d-0.5}^{k+d+0.5} \int_{j-0.5}^{j+0.5} \int_{i-0.5}^{i+0.5} w(x, y, z) F(x, y, z) \mathrm{d}x \mathrm{d}y \mathrm{d}z, \tag{1}$$

where $w(x, y, z)$ is the weight function, and different weight functions can be set in different application scenarios to obtain more accurate pixel values. If $F(x, y, z)$ is known, the corresponding $P_{i,j,k+d}$ can be obtained from the formula 1. Therefore, how to construct a reasonable $F(x, y, z)$ has become the focus of this section. For the convenience of discussion, assuming that $F(x, y, z)$ is defined on $[0.5, m + 0.5] \times [0.5, n + 0.5] \times [0.5, t + 0.5]$, where it satisfies the formula 1 for each pixel $P_{i,j,k+d}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n$. Since the pixel values are generally integers, the formula 1 is approximately valid and the error is less than 0.5. In addition, the weight function $w(x, y, z)$ is set to 1.
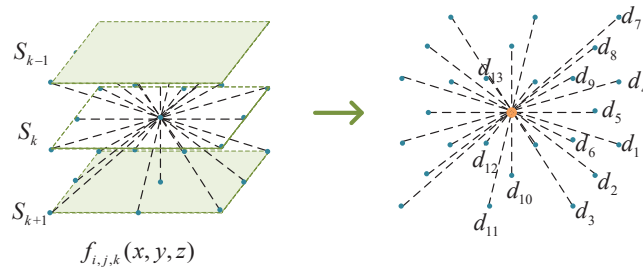
Figure 3: Pixels transform in 13 directions.

3.2 Building the Space Geometry $f_{i,j,k}(x,y,z)$

Since $F(x,y,z)$ can be an arbitrary polynomial function, it is difficult to be constructed accurately only by relying on the original image sequence. Therefore, we construct the approximate space geometry $f_{i,j,k}(x,y,z)$ of the original space geometry. Depending on the numerical approximation theory, any continuous function $T(x,y,z)$ can be expanded into a polynomial Taylor series at a point (x_a, y_a, z_a) in a consecutive three-dimensional space, which is an approximation of $T(x,y,z)$ in the spatial neighborhood of point (x_a, y_a, z_a).

In three-dimensional space, the space geometry $f_{i,j,k}(x,y,z), i=2,3,\ldots,m-1, j=2,3,\ldots,n-1, k=2,3,\ldots,t-1$ must satisfy the following conditions: if $F(x,y,z)$ in formula 1 is a ternary quadratic polynomial, then $f_{i,j,k}(x,y,z)$ can accurately reconstruct $F(x,y,z)$, that is, $f_{i,j,k}(x,y,z)=F(x,y,z)$.

Let $u=x-i, v=y-j, w=z-k$, then $f_{i,j,k}(x,y,z)$ is written on $[-1.5, 1.5] \times [-1.5, 1.5] \times [-1.5, 1.5]$ space of the uvw space geometry:

$$f_{i,j,k}(x,y,z)=a_1 u^2 + a_2 v^2 + a_3 w^2 + a_4 uv + a_5 uw + a_6 vw + a_7 u + a_8 v + a_9 w + a_{10}, \qquad (2)$$

where $a_1, a_2, \ldots, a_{10}$ is unknown coefficients. Next, we will solve these unknown coefficients. In order to reduce the amount of calculation, this paper first discusses how to determine a_7, a_8, a_9, and then calculate the remaining seven coefficients. Here, we take pixel point $\mathsf{P}_{i,j,k}$ as an example for convenient discussion.

Suppose $F(x,y,z)$ can be defined by the formula 2, $\mathsf{P}_{i,j,k}$ is defined by the sampling formula 1, and then it can be inferred that:

$$
\begin{aligned}
&\mathsf{P}_{i,j,k} = \tfrac{1}{12}a_1 + \tfrac{1}{12}a_2 + \tfrac{1}{12}a_3 + a_{10}, \\
&\mathsf{P}_{i-1,j,k} = \tfrac{13}{12}a_1 + \tfrac{1}{12}a_2 + \tfrac{1}{12}a_3 - a_7 + a_{10}, \\
&\mathsf{P}_{i+1,j,k} = \tfrac{13}{12}a_1 + \tfrac{1}{12}a_2 + \tfrac{1}{12}a_3 + a_7 + a_{10}, \\
&\vdots \\
&\mathsf{P}_{i-1,j-1,k-1} = \tfrac{13}{12}a_1 + \tfrac{13}{12}a_2 + \tfrac{13}{12}a_3 + a_4 + a_5 + a_6 - a_7 - a_8 - a_9 + a_{10}, \\
&\mathsf{P}_{i+1,j+1,k+1} = \tfrac{13}{12}a_1 + \tfrac{13}{12}a_2 + \tfrac{13}{12}a_3 + a_4 + a_5 + a_6 + a_7 + a_8 + a_9 + a_{10}.
\end{aligned}
\qquad (3)
$$

The calculated results are as follows:

$$a_7 = e_1 = \left(\mathsf{P}_{i+1,j,k} - \mathsf{P}_{i-1,j,k}\right)/2$$
$$a_8 = e_2 = \left(\mathsf{P}_{i,j+1,k} - \mathsf{P}_{i,j-1,k}\right)/2$$
$$a_9 = e_3 = \left(\mathsf{P}_{i,j,k+1} - \mathsf{P}_{i,j,k-1}\right)/2$$
$$a_7 + a_8 = e_4 = \left(\mathsf{P}_{i+1,j+1,k} - \mathsf{P}_{i-1,j-1,k}\right)/2$$
$$a_7 - a_8 = e_5 = \left(\mathsf{P}_{i+1,j-1,k} - \mathsf{P}_{i-1,j+1,k}\right)/2$$
$$a_7 + a_9 = e_6 = \left(\mathsf{P}_{i+1,j,k+1} - \mathsf{P}_{i-1,j,k-1}\right)/2$$
$$a_7 - a_9 = e_7 = \left(\mathsf{P}_{i+1,j,k-1} - \mathsf{P}_{i-1,j,k+1}\right)/2 \tag{4}$$
$$a_8 + a_9 = e_8 = \left(\mathsf{P}_{i,j+1,k+1} - \mathsf{P}_{i,j-1,k-1}\right)/2$$
$$a_8 - a_9 = e_9 = \left(\mathsf{P}_{i,j+1,k-1} - \mathsf{P}_{i,j-1,k+1}\right)/2$$
$$a_7 - a_8 - a_9 = e_{10} = \left(\mathsf{P}_{i+1,j-1,k-1} - \mathsf{P}_{i-1,j+1,k+1}\right)/2$$
$$a_7 - a_8 + a_9 = e_{11} = \left(\mathsf{P}_{i+1,j-1,k+1} - \mathsf{P}_{i-1,j+1,k-1}\right)/2$$
$$a_7 + a_8 - a_9 = e_{12} = \left(\mathsf{P}_{i+1,j+1,k-1} - \mathsf{P}_{i-1,j-1,k+1}\right)/2$$
$$a_7 + a_8 + a_9 = e_{13} = \left(\mathsf{P}_{i+1,j+1,k+1} - \mathsf{P}_{i-1,j-1,k-1}\right)/2$$

The reconstruction quality of the image edge has the greatest impact on the intuitive visual effect. Therefore, it is necessary to ensure that each surface of the three-dimensional (3D) spatial geometry may reflect the edge characteristics of interpolation image as much as possible. In 3D space, each pixel has some correlation with its surrounding adjacent pixels (up and down, left and right, front and back, diagonal). As shown in Figure 3, in thirteen directions formed in the center pixel and its neighboring pixels, if the image pixel changes slowly in a certain direction, it is more likely that the direction is the image edge, and if the image pixel varies linearly in a certain direction, it is desirable to be a linear function along the direction.

There are thirteen equations in the formula 4. In order to reflect the edge characteristics of the image as much as possible, the constrained least square method is used to determine the three unknown coefficients a_7, a_8, a_9 in this formula. The objective function $G(a_7, a_8, a_9)$ is:

$$G(a_7, a_8, a_9) = w_1(a_7 - e_1)^2 + w_2(a_8 - e_2)^2 + \ldots + w_{13}(a_7 + a_8 + a_9 - e_{13})^2, \tag{5}$$

where $w_i(i = 1, 2, 3, \ldots, 13)$ is the weight function and obtained by addition and subtraction of the equations in the formula 3:

$$\Delta_1 = \mathsf{P}_{i+1,j,k} - 2\mathsf{P}_{i,j,k} + \mathsf{P}_{i-1,j,k} = 2a_1,$$
$$\Delta_2 = \mathsf{P}_{i,j+1,k} - 2\mathsf{P}_{i,j,k} + \mathsf{P}_{i,j-1,k} = 2a_2,$$
$$\Delta_3 = \mathsf{P}_{i,j,k+1} - 2\mathsf{P}_{i,j,k} + \mathsf{P}_{i,j,k-1} = 2a_3,$$
$$\Delta_4 = \mathsf{P}_{i+1,j+1,k} - 2\mathsf{P}_{i,j,k} + \mathsf{P}_{i-1,j-1,k} = 2(a_1 + a_2 + a_4), \tag{6}$$
$$\vdots$$
$$\Delta_{12} = \mathsf{P}_{i+1,j+1,k-1} - 2\mathsf{P}_{i,j,k} + \mathsf{P}_{i-1,j-1,k+1} = 2(a_1 + a_2 + a_3 + a_4 - a_5 - a_6),$$
$$\Delta_{13} = \mathsf{P}_{i+1,j+1,k+1} - 2\mathsf{P}_{i,j,k} + \mathsf{P}_{i-1,j-1,k-1} = 2(a_1 + a_2 + a_3 + a_4 + a_5 + a_6).$$

If $f_{i,j,k}(x, y, z)$ in formula 2 is linearly changed in z axis direction, a_9 should be determined by e_3, that is, w_3 corresponds to a relatively large value. When the center point pixel is close to the linear change in z axis direction, then $\Delta_3 = 2a_3 \approx 0$, thus w_3 is inversely proportional to Δ_3. In the same way, the remaining weight

function can be determined. The weight function is defined as follows:

$$w_i = \frac{\beta}{1 + \alpha \Delta_i^2}, i = 1, 2, \ldots, 13, \tag{7}$$

where α and β are the correction parameters, here the sum of these two parameters is 1. a_7, a_8, a_9 can be obtained by minimizing the objective function:

$$\frac{\partial G(a_7, a_8, a_9)}{\partial a_7} = 0, \ \frac{\partial G(a_7, a_8, a_9)}{\partial a_8} = 0, \ \frac{\partial G(a_7, a_8, a_9)}{\partial a_9} = 0. \tag{8}$$

The remaining parameter $a_1, a_2, \ldots, a_6$ is also determined by the least square method with constraints. Constructing the objective function $G(a_1, a_2, a_3, ..., a_6) = \sum_{\substack{p,q,l=-1,0,1 \\ p \neq q \neq l = 0}} w_{p,q,l}(g_{p,q,l}() - \mathsf{P}_{i+p,j+q,k+l})^2$, where $w_{p,q,l}$ is the weight function, the definition rules are the same as before. If $f_{i,j,k}(x,y,z)$ in formula 2 varies linearly along the z axis, $\mathsf{P}_{i,j,k+1}, \mathsf{P}_{i,j,k}, \mathsf{P}_{i,j,k-1}$ should be on the same line. $\mathsf{P}_{i,j,k+1}$ and $\mathsf{P}_{i,j,k-1}$ should play a decisive role in the change of $f_{i,j,k}(x,y,z)$ along z axis, that is, the considerable value corresponding to $w_{0,0,-1}$ and $w_{0,0,1}$ can be equivalent to w_3 in formula 7. Similarly, other weights can be defined:

$$
\begin{aligned}
&w_{-1,0,0} = w_{1,0,0} = w_1, &&w_{0,-1,0} = w_{0,1,0} = w_2, &&w_{0,0,-1} = w_{0,0,1} = w_3, \\
&w_{1,1,0} = w_{-1,-1,0} = w_4, &&w_{1,-1,0} = w_{-1,1,0} = w_5, &&w_{1,0,1} = w_{-1,0,-1} = w_6, \\
&w_{1,0,-1} = w_{-1,0,1} = w_7, &&w_{0,1,1} = w_{0,-1,-1} = w_8, &&w_{0,1,-1} = w_{0,-1,1} = w_9, \\
&w_{1,-1,-1} = w_{-1,1,1} = w_{10}, &&w_{1,-1,1} = w_{-1,1,-1} = w_{11}, &&w_{1,1,-1} = w_{-1,-1,1} = w_{12}, \\
&w_{1,1,1} = w_{-1,-1,-1} = w_{13}
\end{aligned}
$$

3.3 Correcting Spatial Geometry $f_{i,j,k}(u,v,w)$

In order to ensure that the sampled pixel $\mathsf{V}_{i,j,k+d}$ satisfies the condition $0 \le \mathsf{V}_{i,j,k+d} \le 255$, in three-dimensional space $\Omega_{i,j,k}, \Omega_{i,j,k} = [i-1, j-1, k-1] \times [i+1, j+1, k+1]$, $f_{i,j,k}(u,v,w)$ needs to satisfy $0 \le f_{i,j,k}(u,v,w) \le 255$, which makes the spatial geometry $f_{i,j,k}(u,v,w)$ generated by $f(u,v,w)$ weighted average also satisfy $0 \le f(u,v,w) \le 255$.

Therefore, $f_{i,j,k}(u,v,w)$ can be modified to $f_{i,j,k}(u,v,w) = 255 \times (f_{i,j,k}(u,v,w) - f_{i,j,k}(u_p, v_p, w_p)) \div f_{i,j,k}(u_c, v_c, w_c)$ when the following two cases are true. (1) $f_{i,j,k}(u,v,w)$ reaches the minimum value $f_{i,j,k}(u_p, v_p, w_p) < 0$ at point (x_p, y_p, z_p); (2) $f_{i,j,k}(u,v,w)$ reaches the maximum value $f_{i,j,k}(u_c, v_c, w_c) > 255$ at point (x_c, y_c, z_c).

3.4 Construct Approximate Space Geometry $f(x,y,z)$

This section will discuss how to use $f_{i,j,k}(x,y,z), i = 2, 3, \ldots, m-1, j = 2, 3, \ldots, n-1, k = 2, 3, \ldots, t-1$ to determine the approximation space geometry $f(x,y,z)$.

Specifically, in each area $[i, i+1] \times [j, j+1] \times [k, k+1], i = 1, 2, \ldots, m-1, j = 1, 2, \ldots, n-1, k = 1, 2, \ldots, t-1$, we construct a unit space geometry $B_{i,j,k}(x,y,z)$, $f(x,y,z)$ is then calculated from all $B_{i,j,k}(x,y,z)$ weighted combinations. Here is a discussion about how to construct H in two cases. In the first case, when $i = 2, 3, \ldots, m-1, j = 2, 3, \ldots, n-1, k = 2, 3, \ldots, t-1$, $B_{i,j,k}(x,y,z)$ on the spatial domain $[i, i+1] \times [j, j+1] \times [k, k+1]$ is generated by the $f_{i,j,k}(x,y,z), f_{i+1,j,k}(x,y,z), \ldots, f_{i+1,j+1,k+1}(x,y,z)$ weighted average, that is,

$$B_{i,j,k}(x,y,z) = w_{i,j,k}(x,y,z)f_{i,j,k}(x,y,z) + \cdots + w_{i,j+1,k+1}(x,y,z)f_{i+1,j+1,k+1}(x,y,z), \tag{9}$$

where,

$$w_{i,j,k}(x,y,z) = (1-u)(1-v)(1-w) \quad w_{i+1,j,k}(x,y,z) = u(1-v)(1-w)$$
$$w_{i,j+1,k}(x,y,z) = (1-u)v(1-w) \quad w_{i,j,k+1}(x,y,z) = (1-u)(1-v)w$$
$$w_{i+1,j+1,k}(x,y,z) = uv(1-w) \quad w_{i+1,j,k+1}(x,y,z) = u(1-v)w$$
$$w_{i,j+1,k+1}(x,y,z) = (1-u)vw \quad w_{i+1,j+1,k+1}(x,y,z) = uvw.$$

In the second case, we construct unit space geometry $B_{i,j,k}(x,y,z)$ in the boundary area as follows,

$$\begin{cases} B_{1,j,k}(x,y,z), B_{m-1,j,k}(x,y,z), j=1,\ldots,n-1, k=1,\ldots,t-1 \\ B_{i,1,k}(x,y,z), B_{i,n-1,k}(x,y,z), i=2,\ldots,m-2, k=1,\ldots,t-1 \\ B_{i,j,1}(x,y,z), B_{i,j,t-1}(x,y,z), i=2,\ldots,m-2, j=2,\ldots,n-2. \end{cases}$$

Taking $B_{1,j,k}(x,y,z)$ as an example, as shown in Figure 4, $B_{1,1,1}(x,y,z)$ is defined by $f_{2,2,2}(x,y,z)$, and $B_{1,1,k}(x,y,z)$ is defined by two spatial geometry $f_{i,j,k}(x,y,z)$ when $i=1, j=1, k=2,3,\ldots,t-1$, i.e.,

$$B_{1,1,k}(x,y,z) = f_{2,2,k}(x,y,z)(1-w) + f_{2,2,k+1}(x,y,z)w \tag{10}$$

Then $B_{1,1,k}(x,y,z)$ is defined by four spatial geometry $f_{i,j,k}(x,y,z)$ when $i=1, j=2,3,\ldots,n-1, k=2,3,\ldots,t-1$, i.e.,

$$\begin{aligned} B_{1,1,k}(x,y,z) = &f_{2,j,k}(x,y,z)(1-v)(1-w) + f_{2,j+1,k}(x,y,z)v(1-w)+ \\ &f_{2,j,k+1}(x,y,z)(1-v)w + f_{2,j+1,k+1}(x,y,z)vw. \end{aligned} \tag{11}$$

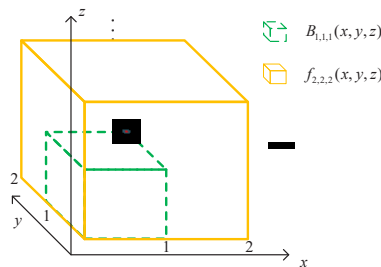Other situations are also followed by the above rules benefited from its symmetry.

Figure 4: $B_{1,1,1}(x,y,z)$ is defined by $f_{2,2,2}(x,y,z)$.

4 EXPERIMENTS AND RESULTS

In this section, for the purpose of comprehensively evaluating the performance of the proposed method, we design the related simulation experiments based on CT datasets. The performance of the proposed method is evaluated from three aspects: visual effects, quantitative analysis and time complexity.

4.1 Dataset

In order to verify the effectiveness of the method in this paper, we employ 448 CT scans for simulation experiments, which were provided by the the First Affiliated Hospital of Shandong First Medical University (Shandong Provincial Qianfoshan Hospital). More specifically, we use the human chest (coronal, cross-sectional) and head tomographic datasets. These three types dataset with different layers spacing have been widely used in medical imaging research tasks. Figure 5 provides example images, and Table 1 lists the related imaging information.

The reconstructed inter-layer image lacks real image basis for visual comparison and quantitative evaluation. Therefore, during the experiments, four consecutive tomographic images were randomly selected from the image sequences of group I, II and III. In each group, the first and the last two tomographic images were used to reconstruct the inter-layer image, while the second image was used as the reference image to evaluate the effect of image interpolation between layers.
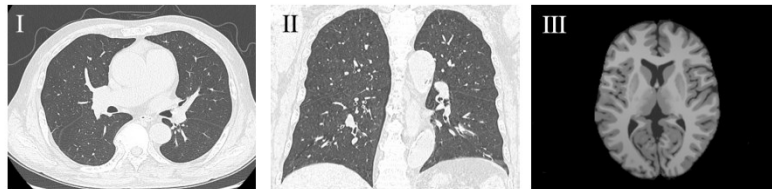

Figure 5: Selected three sets of test image sequences.

Dataset NO.	Data class	Image size	No.image	Layer spacing(mm)
I	Chest(coronal)	560×420	128	2.5
II	Chest(cross-section)	560×420	128	2.5
III	Head	364×436	64	5

Table 1: Experimental Data Information.

4.2 Evaluation Metrics

We compare different approaches via three types of quantitative metrics: average gradient, root mean square error (RMSE) and Structural Dissimilarity (DSSIM)[1].

The average gradient is a non-reference quantitative evaluation method, which represents the average value of all points on an image gradient map. It not only can sensitively reflect the texture change characteristics and small detail differences in images, but also can measure the clarity of the image. The larger the average gradient value, the clearer the image. The calculation formula is as follows:

$$G_{ave} = \frac{1}{M \times N} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left(\frac{\Delta x^2 + \Delta y^2}{2} \right)^{1/2},$$

where $\Delta x = \frac{\partial f}{\partial x}$ represents the gradient along x direction, $\Delta y = \frac{\partial f}{\partial y}$ represents the gradient along y direction, and $M \times N$ represents the image resolution.

RMSE is the root mean square error between the reference image I_{GT} and the estimated interpolation image I_{ip}. The calculation formula is as follows:

$$\mathsf{RMSE} = \left(\frac{1}{M \times N} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left(I_{\mathsf{ip}}(i,j) - I_{\mathsf{GT}}(i,j) \right)^2 \right)^{\frac{1}{2}},$$

where $M \times N$ is the image size. RMSE can be used to measure the pixel error between the interpolation image and the reference image. The smaller the RMSE value, the more ideal the interpolation result, and the higher the accuracy. It is worth noting that, RMSE is more sensitive to outliers, that is, if there is a large difference between the predicted value and the simulated value, RMSE will also be large.

DSSIM is a common indicator to measure the degree of similarity between images, which can better estimate local differences. It is a distance metric based on structural similarity (SSIM). The smaller the DSSIM value, the smaller the degree of image distortion, which is closer to the real image. The calculation formula is as follows:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

$$\mathsf{DSSIM}(x,y) = \frac{1 - \mathrm{SSIM}(x,y)}{2},$$

where μ_x and μ_y represent the gray mean value of the reference image and the interpolation image respectively; σ_x and σ_y are the standard deviations; σ_{xy} is the covariance of x and y; $C_1 = (k_1 L)^2, C_2 = (k_2 L)^2$, and $k_1 = 0.01$, $k_2 = 0.03$, $L = 255$.

4.3 Comparison of Visualization Results

The designated layer is selected from three sets of medical image sequences, and the original image of this layer is simulated as an undistorted inter-layer slice. The designated layer is reconstructed during image quality evaluation and compared with the original image. The actual layer spacing of these three experimental datasets used in the simulation experiment is doubled on the basis of the original layer spacing. Therefore, the interpolation results reconstructed by consecutive tomographic images with one image interval are slightly different from the original image, which is in line with objective reality.
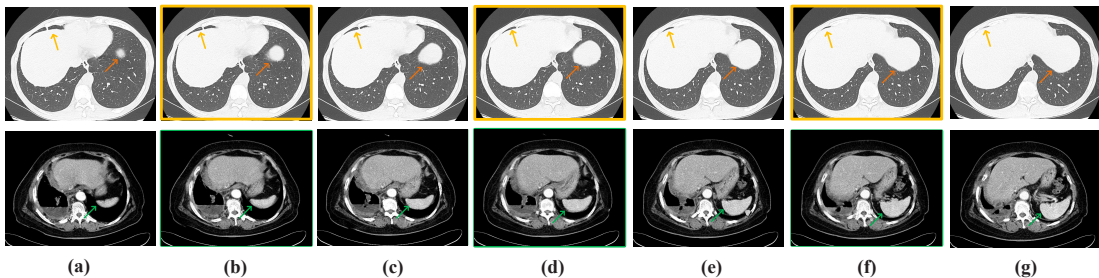

Figure 6: Interpolation results of tomographic medical images: (a) Input image 1; (b) Interpolation result 1; (c) Input image 2; (d) Interpolation result 2; (e) Input image 3; (f) Interpolation result 3; (g) Input image 4.

Two groups of interpolation results are shown in Figure 6 by selecting consecutive CT tomographic images from Chest Dataset (lungs and abdomen). The color rectangular boxes (orange and blue) represent the four
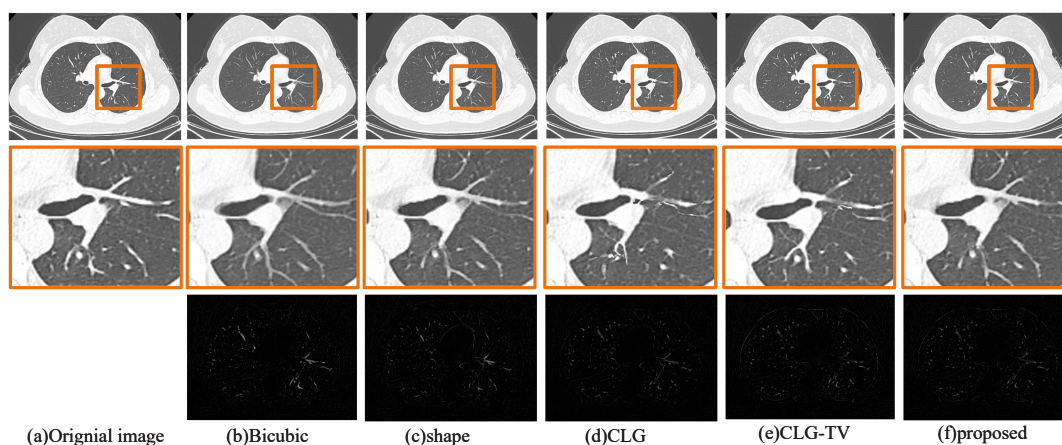
Figure 7: Comparison of interpolation results by different algorithms.

consecutive inputting slices, and the inside of the rectangular box show the inter-lyaer slices reconstructed by the method in this paper. The arrows point to the area of obvious change. From an intuitive point of view, the changes in the contours of the reconstructed images between layers are natural and consistent. The lung CT slice contains information such as lung nodules and lung organ tissue texture. The interpolation results show that the intermediate image obtained by this method does not lose relevant information. The internal details are clear, and the effect of smooth transition can be achieved, which can effectively improve the axis resolution between layers. It is worth noting that interpolation on CT sequence images does not generate new lesion information, but the resolution of CT images is improved to make subsequent three-dimensional reconstruction processing more convenient.

The traditional Bicubic interpolation method [10], Shape-based, CLG [3], CLG-TV [14] and the proposed method were compared in experiments. These four comparison methods are all related classic methods that have been cited and compared in literatures. In terms of visual evaluation, Figure 7, Figure 8, and Figure 9 respectively show the qualitative evaluation results of the above five methods on three different medical image datasets, and the representative visualization results. The rectangular frame is the main change area of the image. By comparing the reference image and the interpolation result, it is obvious that this method can interpolate to obtain a complete inter-layer image in subjective vision, and the edge blur and artifacts of tissues and organs are significantly improved. Compared with the interpolation method based on optical flow registration, the proposed method can achieve better or similar interpolation effect.

In order to compare the subtle differences of the five interpolation results more clearly, the partial magnified diagrams in color rectangular boxes are shown in Figure 7, Figure 8, and Figure 9, respectively. By enlarging the local details of the image, the quality of the image generated by various interpolation methods can be more intuitively evaluated from the visual effects. It can be seen from thees figures that bicubic interpolation only uses the voxel center of the slice to be interpolated as the interpolation point, and is simply calculated by using the weighted average of the 16 nearest sampling points. It cannot accurately reflect the non-rigid deformation process. The edges are obviously blurred. The shape-based method has significantly improved the artifacts in the reconstruction results, but the quality of the interpolated image is not high in the case of more deformation details.

Based on CLG method, CLG-TV method uses complementary information between different phases to enhance the structure and texture details of the reconstructed image, which can reflect the non-rigid deformation process, effectively eliminate the edge artifacts, and retain the contour information. But the phenomenon
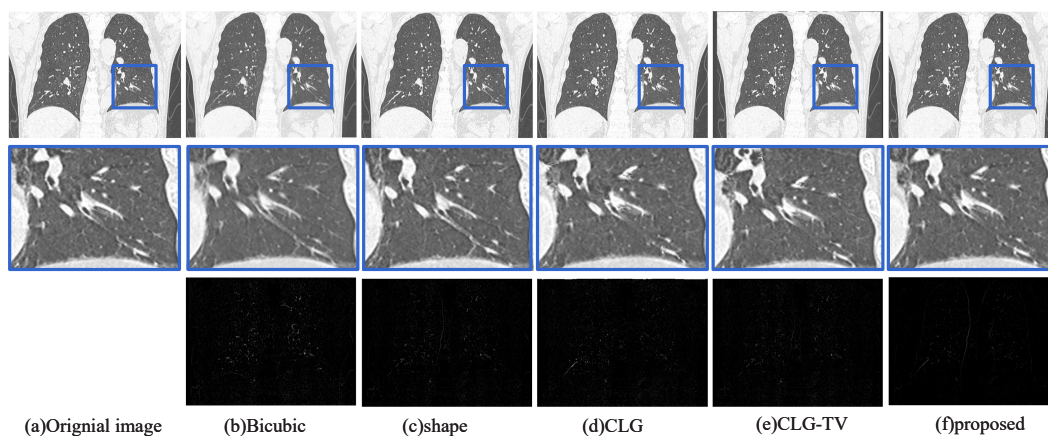
Figure 8: Comparison of interpolation results of various algorithms 2.
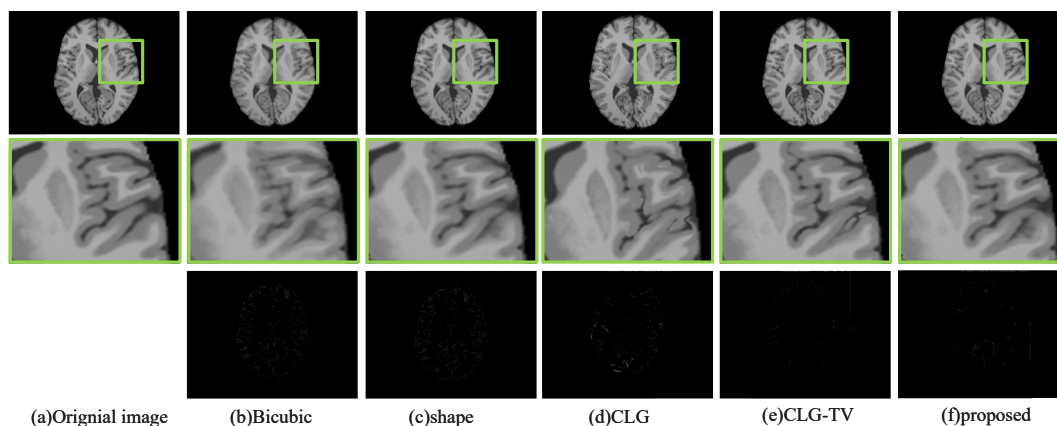

Figure 9: Comparison of interpolation results of various algorithms 3.

of tiny holes undountly affects the quality of the interpolated image. The method in this paper samples three consecutive images to construct a 3D space geometry, because it considers the correlation effect of 108 neighboring points in the area around the point to be interpolated, and further eliminates the problem of gray value discontinuity based on gray-level interpolation. It can solve the phenomenon of edge blur appeared in the bicubic interpolation, reduce the difference between the obtained interpolation result and the original slice significantly.

In order to compare them more intuitively, Figure 7, Figure 8, and Figure 9 show the error image comparison between these five methods and the original inter-layer image. It can be seen intuitively from the difference map that this method has smaller error with the real image, which further shows that it can accurately reconstruct CT images. By observation, this method is closer to the real image than the other four methods in the area where the deformation is smaller.

Furthermore, 3D reconstruction experiments were carried out using consecutive lung CT images with 2.5 mm layer spacing provided by the First Affiliated Hospital of Shandong First Medical University. We use this method to reconstruct the inter-layer slices, then model the virtual lung tissues and organs via the medical

image visualization platform 3D Slicer. The effect is shown in Figure 10. It can be seen that the direct 3D reconstruction of lung tomographic images is not good enough, with obvious step-like phenomena and rough tissue boundaries or even fractures. After reconstructing the inter-layer slice by this method to reduce the layer spacing to 1.25mm, the 3D reconstruction effect is relatively perfect, and the step-like phenomenon disappears obviously. Further, after 3D reconstruction with 0.625mm layer spacing, the step-like phenomenon basically disappeared, the details and texture of blood vessels in the lung were greatly improved, which conformed to the structural characteristics of human organs. The result can meet the requirements of 3D reconstruction and visualization of medical images.
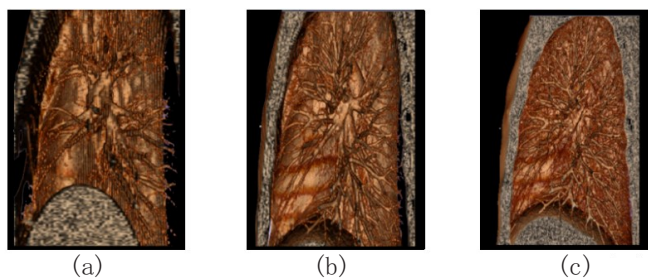

Figure 10: 3D reconstruction model of lung images: (a) The layer spacing is 2.5mm; (b) The layer spacing is 1.25mm; (c) The layer spacing is 0.625mm.

4.4 Quantitative Evaluations

In order to further illustrate the feasibility of the method in this paper, Table 2 shows the quantitative data of the reconstruction results from different methods, where the rows represent RMSE and DSSIM from top to bottom. Combined with the data in Table 2, it can be seen that, the method proposed in this paper has the lowest average RMSE and DSSIM values, and has a greater advantage in objective numerical evaluation. Compared with the bicubic interpolation, DSSIM value of this method is reduced by about 46.59% on average. It is reduced by 18.55%, 15.13% and 5.77% on average compared to Shape-based interpolation, CLG method and CLG-TV interpolation method respectively.

By analyzing the quantitative results of three different datasets, it is found that this method performs better on the first and second image sequences than the third group. It can be seen that this method can better interpolate the interval between the layers. The image is less obvious for larger layer spacing, similar to the CLG-TV method. The results show that the performance of this method in improving the spatial resolution of medical images is better than bicubic interpolation, Shape-based method and CLG method, and slightly better or equivalent to CLG-TV method.

The change rate of the gray value near the edge of the image can reflect the clarity of the image. In order to further illustrate that the method in this paper can effectively improve the edge blur, the objective index of the average gradient is used to measure the clarity of the method. From Figure 11, the reconstruction result obtained by this method is higher than the average gradient value obtained by the linear and shape-based methods, and the image definition is obviously enhanced, which is similar to the result obtained by the CLG-TV method to a certain extent. Our goal is to synthesize the pixels of the interpolated image while making the interpolation result appear real and smooth in the human eyes.

As we all know, the existing quantitative evaluation results are not a good substitute for human visual perception, and the continuity in time cannot be measured by numerical standards. We found that the visual quality and competitiveness of our method are far greater than the numerical difference. The proposed method

Image Dataset	Metrics	Bicubic	Shape-based	CLG	CLG-TV	Proposed
I	RMSE	19.3265	14.7652	16.1897	12.8430	11.4038
	DSSIM	0.2579	0.2105	0.1891	0.1773	0.1577
II	RMSE	8.2821	7.1558	6.7615	5.7551	3.8255
	DSSIM	0.0948	0.0205	0.0237	0.0142	0.0128
III	RMSE	21.4194	21.6622	21.8338	21.7167	20.5492
	DSSIM	0.1701	0.1120	0.1162	0.1048	0.1089
Average	RMSE	16.3427	14.5277	14.9283	13.4384	11.9262
	DSSIM	0.1743	0.1143	0.1097	0.0988	0.0931

Table 2: Comparison of RMSE and DSSIM values of different interpolation methods.

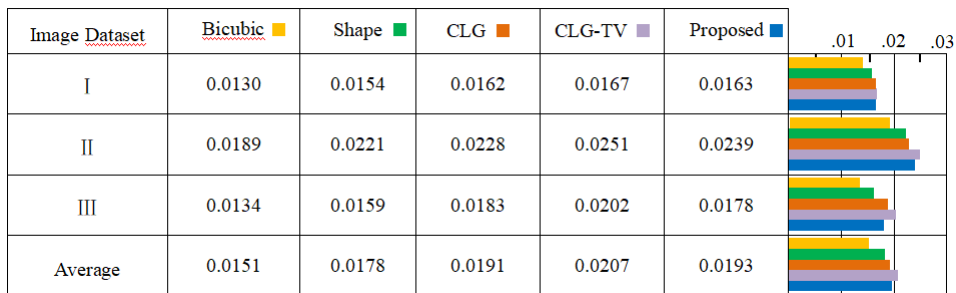| Image Dataset	Bicubic ▪	Shape ▪	CLG ▪	CLG-TV ▪	Proposed ▪	
I	0.0130	0.0154	0.0162	0.0167	0.0163	
II	0.0189	0.0221	0.0228	0.0251	0.0239	
III	0.0134	0.0159	0.0183	0.0202	0.0178	
Average	0.0151	0.0178	0.0191	0.0207	0.0193	

Figure 11: Comparison of average gradients of different methods.

uses the priori information of upper and lower layers to reconstruct the inter-layer images, so no new lesion information which could be misleading to medical staff is generated. In summary, the contour artifacts in the interpolation result of the proposed method are reduced. Furthermore, the intermediate state of the smooth transition of the organ tissue structure from one image to another can be simulated, which is beneficial to clinical diagnosis. As shown in Figure 10, the results of 3D reconstruction have greatly improved the internal details of organs which medical staff could otherwise have missed.

In actual clinical application, the system needs not only higher image quality, but also the real-time calculation speed. The faster the calculation speed, the higher the practicability of the system. In the process of solving the polynomial coefficients of the proposed method, assuming that the number of pixels contained in the experimental image is $m \times n$, the time complexity of the process is $O((m - 2) \times (n - 2))$. The time complexity of the process of synthesizing pixels is $O(m \times n)$, so the total time complexity is: $\max \{O((m - 2) \times (n - 2)), O(m \times n)\}$, that is, $O(mn)$.

Running on the Intel(R) Core(TM) i7-4790 CPU @ 3.60 GHz, 16 GB memory, the time required to reconstruct a single image on three different datasets by five methods evaluated is as shown in Table 3. It can be seen that this method greatly improves the time efficiency and reduces the reconstruction complexity of a single image while obtaining the same accuracy as CLG and the CLG-TV methods. It is close to the time efficiency of the bicubic interpolation method, but the effect of the bicubic interpolation is not as good as the method in this article in both qualitative and quantitative. Therefore, the method in this paper not only achieves better reconstruction results in visual evaluation and objective values, but also has low time

Image Dataset	Bicubic	Shape	CLG	CLG-TV	Proposed
I	1.261	61.574	76.112	68.145	**0.834**
II	**0.437**	48.957	72.374	62.309	1.331
III	**0.322**	40.295	73.260	59.628	1.035
Average	**0.673**	50.275	73.915	63.361	1.067

Table 3: Comparison of running time (second) of different interpolation methods.

complexity.

5 CONCLUSION

This paper proposes a slice interpolation method for medical image based on spatial geometry polynomial fitting. This method uses three consecutive CT slices in 3D space to reversely reconstruct the fitting space geometry that approximates the original space scene, and performs the calculation on the space geometry. Resampling strategy helps to generate new voxels and reconstruct inter-layer slices. We carefully evaluate our approach on three different CT datasets and find that, comparing with other state-of-the-art methods, the proposed method produces improvement in terms of visual quality and quantitative measures. More importantly, the proposed method improves the efficiency of slice interpolation for medical image. At the same time, after interpolating the inter-layer slice via this method, the layer spacing of the medical CT image is reduced. On this basis, 3D reconstruction is performed, the step-like phenomenon basically disappears, the details and texture of the blood vessels in the lung are greatly improved, which is in line with human organs. The structural features can meet the requirements of 3D reconstruction and visualization of medical images.

However, the sampling point range of this method is limited to 108 pixels around the area, where the prior information is relatively simple. As a result, when the anatomical structure on two consecutive slices varies greatly, the interpolation effect has limitations, such as artifacts or motion-blurriness. In the future, we will expand the range of sampling points to fully describe the semantic information of the points to be interpolated, and conduct more experimental demonstrations on different scenarios and types of data. And the future work also includes researching new functions to construct space geometry to improve the interpolation accuracy.

ACKNOWLEDGEMENTS

REFERENCES

[1] Akl, A.; Yaacoub, C.: Image analysis by structural dissimilarity estimation. In 2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA), 1–4. IEEE, 2019. http://doi.org/10.1109/ipta.2019.8936086.

[2] Baghaie, A.; Yu, Z.: Curvature-based registration for slice interpolation of medical images. In International Symposium Computational Modeling of Objects Represented in Images, 69–80. Springer, 2014. http://doi.org/10.1007/978-3-319-09994-1_7.

[3] Bruhn, A.; Weickert, J.; Schnörr, C.: Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. International journal of computer vision, 61(3), 211–231, 2005. http://doi.org/10.1023/b:visi.0000045324.43199.43.

[4] Goshtasby, A.; Turner, D.A.; Ackerman, L.V.: Matching of tomographic slices for interpolation. IEEE Transactions on Medical Imaging, 11(4), 507–516, 1992. http://doi.org/10.1109/42.192686.

[5] Grevera, G.J.; Udupa, J.K.: Shape-based interpolation of multidimensional grey-level images. IEEE transactions on medical imaging, 15(6), 881–892, 1996. http://doi.org/10.1109/42.544506.

[6] Guo, Y.; Bi, L.; Ahn, E.; Feng, D.; Wang, Q.; Kim, J.: A spatiotemporal volumetric interpolation network for 4d dynamic medical image. IEEE, 2020. http://doi.org/10.1109/cvpr42600.2020.00478.

[7] Hou, H.; Andrews, H.: Cubic splines for image interpolation and digital filtering. IEEE Transactions on acoustics, speech, and signal processing, 26(6), 508–517, 1978. http://doi.org/10.1109/tassp.1978.1163154.

[8] Hu, Y.; Li, Y.; Song, R.: Robust interpolation of correspondences for large displacement optical flow. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 481–489, 2017. http://doi.org/10.1109/cvpr.2017.509.

[9] Huang, Q.; Huang, Y.; Hu, W.; Li, X.: Bezier interpolation for 3-d freehand ultrasound. IEEE Transactions on Human-Machine Systems, 45(3), 385–392, 2014. http://doi.org/10.1109/THMS.2014.2374551.

[10] Keys, R.G.: Cubic convolution interpolation for digital image processing. IEEE Transactions on Acoustics, Speech, and Signal Processing, 29, 2003. http://doi.org/10.1109/TASSP.1981.1163711.

[11] Kudo, A.; Kitamura, Y.; Li, Y.; Iizuka, S.; Simo-Serra, E.: Virtual thin slice: 3d conditional gan-based super-resolution for ct slice interval. In International Workshop on Machine Learning for Medical Image Reconstruction, 91–100. Springer, 2019. http://doi.org/10.1007/978-3-030-33843-5_9.

[12] Laghrib, A.; Hadri, A.; Hakim, A.; Raghay, S.: A new multiframe super-resolution based on nonlinear registration and a spatially weighted regularization. Information Sciences, 493, 34–56, 2019. http://doi.org/10.1016/j.ins.2019.04.029.

[13] Leng, J.; Xu, G.; Zhang, Y.: Medical image interpolation based on multi-resolution registration. Computers & Mathematics with Applications, 66(1), 1–18, 2013. http://doi.org/10.1016/j.camwa.2013.04.026.

[14] Liu, H.; Lin, Y.; Ibragimov, B.; Zhang, C.: Low dose 4d-ct super-resolution reconstruction via inter-plane motion estimation based on optical flow. Biomedical Signal Processing and Control, 62, 102085, 2020. http://doi.org/10.1016/j.bspc.2020.102085.

[15] Liu, X.; Chen, L.; Wang, W.; Zhao, J.: Robust multi-frame super-resolution based on spatially weighted half-quadratic estimation and adaptive btv regularization. IEEE Transactions on Image Processing, 27(10), 4971–4986, 2018. http://doi.org/10.1109/TIP.2018.2848113.

[16] Peng, C.; Lin, W.A.; Liao, H.; Chellappa, R.; Zhou, S.K.: Saint: spatially aware interpolation network for medical slice synthesis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 7750–7759, 2020. http://doi.org/10.1109/cvpr42600.2020.00777.

[17] Penney, G.P.; Schnabel, J.A.; Rueckert, D.; Viergever, M.A.; Niessen, W.J.: Registration-based interpolation. IEEE Transactions on Medical Imaging, 23(7), 922–926, 2004. http://doi.org/10.1109/TMI.2004.828352.

[18] Raya, S.P.; Udupa, J.K.: Shape-based interpolation of multidimensional objects. IEEE transactions on medical imaging, 9(1), 32–42, 1990. http://doi.org/10.1109/42.52980.

[19] Rueckert, D.; Aljabar, P.; Heckemann, R.A.; Hajnal, J.V.; Hammers, A.: Diffeomorphic registration using b-splines. In International Conference on Medical Image Computing and Computer-Assisted Intervention, 702–709. Springer, 2006. http://doi.org/10.1007/11866763_86.

[20] Santarelli, C.; Argenti, F.; Uccheddu, F.; Alparone, L.; Carfagni, M.: Volumetric interpolation of tomographic sequences for accurate 3d reconstruction of anatomical parts. Computer methods and programs in biomedicine, 194, 105525, 2020. http://doi.org/10.1016/j.cmpb.2020.105525.

[21] Wang, Z.; Liu, J.; Chen, X.; Li, G.; Han, H.: Sparse self-attention aggregation networks for neural sequence slice interpolation. BioData Mining, 14(1), 1–19, 2021. http://doi.org/10.1186/s13040-021-00236-z.

[22] Wu, Z.; Wei, J.; Yuan, W.; Wang, J.; Tasdizen, T.: Inter-slice image augmentation based on frame interpolation for boosting medical image segmentation accuracy. arXiv preprint arXiv:2001.11698, 2020.

[23] Zulkifli, N.A.B.; Karim, S.A.A.; Sarfraz, M.; Ghaffar, A.; Nisar, K.S.; et al.: Image interpolation using a rational bi-cubic ball. Mathematics, 7(11), 1045, 2019. http://doi.org/10.3390/math7111045.