Anisotropic Phase Stretch Transform-based Algorithm for Segmentation of Activated Sludge Phase-contrast Microscopic Image

PENGFEI XU¹, ZHIQING ZHOU¹, HESHENG SHI¹, and ZEXUN GENG¹,²

¹Department of Information Engineering, Pingdingshan University, Pingdingshan 467000, CHN
²Information Engineering University, Zhengzhou 450052, CHN
Corresponding author: Pengfei Xu (e-mail: 3221@pdsu.edu.cn).

This work was supported by the Key Scientific Research Projects of Henan Higher Education Institutions under Grant 21A520034.

ABSTRACT The activated sludge (AS) process is a biological treatment process of wastewater used in sewage treatment plants, in which settling of AS is vitally important for wastewater treatment. In AS, however, bulking caused by filamentous bacteria will significantly reduce the settling capacity of activated sludge. Traditionally, the physicochemical method has been used to monitor the status or performance of filamentous bacteria in activated sludge, while it is a very compromising means in modern digital quality control for wastewater treatment when digital image processing and analysis technology is used to determine the status of filamentous bacteria in activated sludge. In order to avoid the disadvantages of deficient direction selectivity in the translation vector field for isotropic phase stretch transform (PST) and high noise susceptibility in determining the status of filamentous bacteria in AS and low accuracy, an anisotropic PST (APST) method for filamentous bacteria test in AS was proposed in this paper; specifically speaking, by analyzing and deriving the disadvantages of the traditional isotropic PST kernel function, we designed an APST kernel function, put forward a non-maximum suppression strategy for processing of images from different aspects and combined with the relative total variation theory to form APST-based algorithm for segmentation of activated sludge phase-contrast microscopic image. According to a large number of experimental results, in terms of the overall segmentation effect or the subjective and objective evaluation indicators, the resultant algorithm in this paper is superior to the latest traditional PST segmentation algorithm, Canny image segmentation algorithm and Sobel image segmentation algorithm. Also, the segmentation results for filamentous bacteria showed that probability rand index (PRI) and global consistency error (GCE) indicators were all improved by about 30%, supporting the effectiveness and superiority of the algorithm in this paper.

INDEX TERMS Activated sludge phase-contrast image, bacteria segmentation, anisotropic phase stretch transform, relative total variation

I. INTRODUCTION Activated sludge is the grain of flocculent structure formed by mixture of microorganisms, suspended substances and colloids etc. featuring strong adsorption performance and favorable settling ability [1]. The stability of floc is the precondition for favorable settling ability, while favorable settling ability is the foundation for stability of processing system. The decreased settling ability of activated sludge can lead to sludge bulking and decreased processing effect, or loss of floc along the water, which will deteriorate the effluent quality [2]. There are multiple factors affecting the settling ability of activated sludge, microscopically including sludge floc morphological structure, flocculation capacity, viscosity, hydrophilia [3] and surface
electrical behavior [4], floc size and distribution characteristics [5], sludge concentration [6], EPS (Extracellular Polymeric Substances) content, ion strength, positive ion concentration, microbial population and their activity, filamentous bacteria type and quantity etc [7]. However, the measurement of the above-mentioned factors involves high cost, complex operation and long duration. With the development of digital image processing and microscopic technology, AS image processing and analysis based on digital image becomes an effective potential means for monitoring the microorganism state and sludge bulking in sewage treatment plants [8].

Image segmentation is a key step in image analysis and the basis for further understanding of the microbial structure. The accuracy of image analysis depends on the quality of the segmentation of microbial aggregates and filaments in microscopy images of activated sludge. Several traditional image segmentation algorithms, including edge detection, clustering, texture-based segmentation, watershed algorithm, and some combinations of algorithms, have been reported for activated sludge images[9]. Khan et al. proposed a robust segmentation procedure for flocs and filamentous bacteria, and investigated regression models for SVI on the basis of the extracted morphological characteristic parameters of filaments and flocs[10]. Shellehammer et al. adopted contemporary classification convolutional networks into fully convolutional networks (FCNs) and transfer their learned representations to the segmentation task[11]. The U-Net deep learning network is a potential tool that can be used to further improve the loss of detailed image information caused by the multiple down-sampling operations of the image in the network[12]. The U-Net model network structure is easily trained and is suitable for a small sample. Li-Jie Zhao et al. proposed segmentation of activated sludge phase contrast microscopy images using U-Net deep learning model[13].

The analysis and processing of phase-contrast image of activated sludge are divided into 5 stages, including image acquisition and display, image enhancement, image segmentation, mathematical morphology processing and feature information extraction. Therein, image segmentation and feature information extraction are to extract the significant features or those to be investigated in the image (e.g., features of filamentous bacteria and flocs), which is the most critical step during the process of image analysis. The image segmentation methods extensively used mainly include threshold segmentation, clustering segmentation, region growing, deep learning, genetic-based segmentation, edge detection-based segmentation method (e.g., Canny-based image segmentation) and phase-based image segmentation, etc. Nisar et al. [14] adopted phase stretch transform (PST) for segmentation of phase-contrast image of activated sludge, which gets a good segmentation. However, in the paper, the main theory of phase warp and stretch transform was not analyzed in depth, and the traditional isotropic phase kernel function was still adopted. PST (phase stretch transform) is the theory and method proposed by M.H. Asghari and B. Jalali[15] in 2015 for digital signal and image processing, and its fundamental principle is to adopt warp and translation of signal phase after Fourier transform so as to highlight the phase information of high-frequency component and enlarge the angle value corresponding to the edge in the image after inverse Fourier transform, thereby acquiring the medium-high frequency features of image (e.g., edge and texture etc.). The kernel function for phase translation initially designed by M.H. Asghari and B. Jalali is an isotropic inverse tangent function without direction selectivity, and shows equal phase translation of different frequency spectrum on the concentric circle of the same radius on the plane for frequency domain (u, v), so that the detected edge contains massive noises, and especially the isolated, finely-divided high-frequency noisy points are reserved as the high-frequency components in the image by mistake, which brings difficulty in the subsequent threshold processing. Through theoretical analysis in this paper, a kind of anisotropic kernel function model for phase translation was proposed, which extended the contents of PST theory put forward by M.H. Asghari and B. Jalali, and the anisotropic PST was applied to the segmentation between filamentous bacteria and floc microorganism in the phase-contrast image of activated sludge and a favorable effect was obtained.

The contents of this paper were arranged as follows: in section 2, we introduced the background and the basic theory of PST, and analyzed the design and features of the typical isotropic kernel functions for phase translation; in addition, we designed an anisotropic PST kernel function for translation; besides, the rationality of the anisotropic kernel function was demonstrated theoretically to put forward a non-maximum suppression strategy for
anisotropic translation vector field. According to the experimental results, the anisotropic kernel function proposed in this paper is not only superior to the traditional isotropic kernel function, but also better than the edge feature detection performance of the typical operators such as Canny and Sobel. In section 3, we compared and analyzed the experimental results of edge feature extraction and segmentation of microscopic image of bacteria. Section 4 describes the conclusion.

II. SEGMENTATION ALGORITHM FOR ACTIVATED SLUDGE BASED ON APST

A. The algorithm flow is as follows:

![Algorithm flow chart](image)

B. RTV filtering processing

Relative Total Variation (RTV) is an image processing method based on local variation measure. Thanks to its accurate measurement of absolute variation and internal variation within the local neighborhood of the image, it can effectively distinguish the fine texture in the image from the main structure. It can not only be used for structure-texture decomposition of image and image editing, but also as an effective filtering tool suitable for reservation of fine structure of image [16].

RTV firstly defines local windowing total variation measure $D_x(p), D_y(p)$ [16] of every pixel p on any image S:

$$D_x(p) = \sum_{q\in R(p)} g_{p,q} \cdot (\partial_x S)_q | D_y(p)$$

R(p) is a local rectangle window centering on pixel p; $q$ belongs to $R(p)$; $g_{p,q}$ is the weighting factor defined according to spatial affinity. Apparently, $D_x(p)$ and $D_y(p)$ measured absolute spatial change of image gray value within local range around pixel p, and then defined the inherent variations $L_x(p)$ and $L_y(p)$ of the local window:

$$L_x(p) = \sum_{q\in R(p)} g_{p,q} \cdot (\partial_x S)_q$$

$$L_y(p) = \sum_{q\in R(p)} g_{p,q} \cdot (\partial_y S)_q$$

It is noticed that $L_x(p)$ and $L_y(p)$ measured overall spatial change of image gray value within local range around pixel p, while RTV measure is expressed as:

$$RTV(p) = \frac{D_x(p)}{L_x(p) + \epsilon} + \frac{D_y(p)}{L_y(p) + \epsilon}$$

The small positive number $\epsilon$ of denominator is used to avoid denominator from being zero.

If pixel p is located on or near the major structure or salient edge, the gradient of each pixel within local range centering on p at two directions shall have approximately identical symbols, or approximate/similar gradient direction, i.e. approximately $(\partial_x S)_q \cdot (\partial_x S)_t > 0$ and $(\partial_y S)_q \cdot (\partial_y S)_t > 0$ are obtained on different pixel points s and t in the neighborhood of p, which enables close values between $D_x(p)$ and $L_x(p)$, between $D_y(p)$ and $L_y(p)$;
however, when pixel p is located in the messy and irregular fine texture region, opposite sign of gradient value at different pixel s and t will occur, which leads to the fact that the local measure $L_s(p)$ centering on p is much smaller than $D_s(p)$ and $L_s(p)$ is much smaller than $D_s(p)$. As a result, it is concluded that when p is located on or near major structure or outline edge, RTV(p) is approximate to 2; when pixel p is located in the messy and irregular fine texture region, RTV(p) is much larger. Therefore, in distinction of significant structure and fine texture, in order to enhance the contrast between texture and the major structure, especially for the visually obvious region or feature, measures L and D should be combined to form a more effective regularized term in structure-texture decomposition, i.e., the second term in equation (4), and the following edge-preserving filtering is formed:

$$\arg\min_p \sum_p (S_p - I_p)^2 + \lambda \left( \frac{D_s(p)}{L_s(p) + \epsilon} + \frac{D_s(p)}{L_s(p) + \epsilon} \right)$$

(4)

Where, I means input image; $S$ is the resulting structure image (i.e., filtering output image) extracted from the input image; $(S_p - I_p)^2$ is the fidelity term (avoid substantial deviation between input and output image); the regularization term RTV(p) (the second term in equation (4)) measured whether local image is the major structure region or fine texture-noise region. A smaller parameter $\lambda$ enables the second term in (4) to maintain large value during the process of minimization, which restrains output image $S$ to reserve major structure and protruding edge (i.e., texture) within the local region near pixel $p$, inhibits fine and messy texture and isolates noise simultaneously to successively realize elimination of tiny texture from input image and reservation of major structures including protruding edge and structural texture, etc.

In this paper, in order to better extract filamentous bacteria and flocs and other microorganisms from the phase-contrast microscopic image of activated sludge, de-noising shall be performed for the input image before feature extraction of image. Though the comparison of feature extraction results after ordinary Gauss low-pass filtering and RTV filtering, it can be seen from Figure 2 and 3 that the effect of de-noising and feature extraction after RTV filtering was significantly superior to that of traditional Gauss low-pass filtering; thus, RTV image filtering was adopted in the preprocessing stage in this paper.
P. Xu et al.: Anisotropic Phase Stretch Transform-based Algorithm for Segmentation of Activated Sludge Phase-contrast Microscopic Image

FIGURE 3. Example 2 of comparison of anisotropic image segmentation between RTV filtering and ordinary low-pass filtering

C. EDGE FEATURE EXTRACTION BASED ON APST

1) INTRODUCTION OF PST

PST, a physics-inspired digital image transformation, which was proposed by M.H. Asghari and B. Jalali in 2015, simulates the process of electromagnetic wave propagation in the diffraction medium with a dielectric function that has warped dispersive (frequency dependent) property. It simulates the diffraction process through specific all-pass phase filter \( H(\omega) = \exp\{j\beta(\omega)\} \) with frequency-dependent diverging, in which phase \( \beta(\omega) \) group delay \( \tau(\omega) = \frac{\partial \beta(\omega)}{\partial \omega} \) is “S”-shape linear or sub-linear function, able to reshape signal field before signal sampling at analog-to-digital conversion to realize compression of analog signal bandwidth without prolonged duration of the signal in time domain, i.e. reducing TBP (time bandwidth product) of the signal to solve the two inherent problems of the traditional Nyquist uniform sampling[15]: first, when sampling rate is given, the traditional Nyquist uniform sampling can only acquire the information which is two times the maximum frequency components of the signal; second, when there is redundant in analog signal, the traditional Nyquist uniform sampling will obtain far more samples than necessary since oversampling occurs for the signals lower than Nyquist frequency).

PST that can be used before analog-to-digital conversion (ADC) can also be applied in the field of digital signal processing. The process of digital image edge detection -based PST was shown in Figure 4, in which the source image was subject to smoothing through a local low-pass filter kernel function, a phase operation of non-linear frequency function called phase stretch (dispersion) transform (PST) in frequency domain, and finally post-processing including thresholding and morphological filter to achieve edge detection.

The mathematical model of PST in frequency domain is as follows [15]

\[
A(m,n) = \angle \text{IFFT}_2 \left\{ \tilde{K}(u,v) \cdot \tilde{L}(u,v) \cdot \text{FFT}_2 \left\{ B(m,n) \right\} \right\}
\]

where, \( A(m,n) \) is the output phase image; “\( \angle \)” is the angle operator; \( B(m,n) \) means the original input image; \( \text{FFT}_2 \) and \( \text{IFFT}_2 \) are 2D fast Fourier transform and inverse transform respectively; \((u,v)\) stands for the variable frequency; \( \tilde{L}(u,v) \) is the frequency response of local low-pass smoothing filter; \( \tilde{K}(u,v) = e^{-j\varphi(u,v)} \) is the non-linear kernel function for phase warp relying on frequency; \( \varphi(u,v) \) is the non-linear function of the frequency variable.

FIGURE 4. Diagrammatic drawing of the image edge detection process of PST
2). BASIC PST KERNEL FUNCTION

Although any phase kernel function can be considered in PST, according to the literature [15], the derivative of kernel function \( \phi(u,v) \), i.e., group delay, is the linear or sub-linear function of the variable frequency. One simple example is “S”-shaped inverse tangent function. For simplicity, assuming that this operation for phase warp is isotropic on the plane of frequency domain, the warp degree is only related to the polar radius \( r \) under the polar coordinate system on \( o-uv \) frequency plane but irrelevant to polar angle \( \theta \); that is, assuming that the kernel function prototype of PST shows a circular symmetry in terms of the variable frequency, \( \phi'(u,v) = \phi'_\text{polar}(r,\theta) = \phi'_\text{polar}(r) \) (6)

Where, \( r \) is the polar radius under the polar coordinate system on \( o-uv \) frequency plane; \( \theta \) is polar angle and its relation with the variable frequency \( uv: r = \sqrt{u^2+v^2} \), \( u = r \cos \theta \), \( v = r \sin \theta \), \( \theta = \tan^{-1}\left(\frac{v}{u}\right) \). Let that \( \phi'_\text{polar}(r) \), the derivative of \( r \) is S-shaped inverse tangent function, then

\[
\frac{d\phi'_\text{polar}}{dr} = \tan^{-1}(r)
\]

It is noticed that the \( uv \) frequency plane after Fourier transform is a finite region, thus, \( \phi'_\text{polar}(r) \) can be solved by equation (7).

\[
\phi'_\text{polar}(r) = \int_0^r \tan^{-1}(x)dx = \int_0^r \frac{x}{1+x^2}dx = r \tan^{-1}r - \frac{1}{2} \ln(1+r^2)
\]

\( \phi_g \) is obtained through normalization of phase function in equation (8):

\[
\phi_g(r) = \frac{r \tan^{-1}r - 1/2 \cdot \ln(1+r^2)}{r_{\text{max}} \tan^{-1}r_{\text{max}} - 1/2 \cdot \ln(1+r_{\text{max}}^2)}
\]

Phase stretch strength \( S \) and warp \( W \) in non-linear warp and stretch transform are added into the phase function in equation (9) to obtain the final kernel function for phase translation \( \phi_g(r,W,S) \) with strength parameter \( S \) and warp parameter \( W \) in PST:

\[
\phi_g(r,W,S) = S \cdot \frac{W r \tan^{-1}(W \cdot r) - 1/2 \cdot \ln(1+(W \cdot r)^2)}{\max(W) \tan^{-1}(\max(W) \cdot r_{\text{max}}) - 1/2 \cdot \ln(1+(\max(W) \cdot r_{\text{max}})^2)}
\]

(10)

Where, \( \tan^{-1}(\cdot) \) means inverse tangent function; \( \ln(\cdot) \) is natural logarithm; \( r_{\text{max}} \) is the maximum frequency. Equation (10) is the phase stretch kernel function used in the literature [15].

Let’s say the Fourier transform of source image \( B(x,y) \) is \( \hat{B}(u,v) \) then \( \hat{B}(u,v) = \hat{B}(u,v) |\phi_g(u,v) \), \( (u,v) \) is frequency variable. When it is substituted into PST image edge extraction model in equation (5) (low-pass filtering is not considered for now, i.e. \( L(u,v) = 1 \)),

\[
A(m,n) = \text{IFFT2}\left\{|\hat{B}(u,v) |\phi_g(u,v) + \phi_g(W,S)\right\}
\]

(11)

It can be seen from equation (11) that PST is essentially to perform a shift \( \phi_g(r,W,S) \) to the phase angle \( \phi_g(u,v) \) of \( \hat{B}(u,v) \), in which the translation \( \phi_g(r,W,S) \) is the increasing function of \( r \), i.e. the translation is big for high-frequency components (high \( r \)). The high-frequency information in the image is highlighted and edge extraction is realized through applying big translation to the high-frequency components of \( B(x,y) \). On the frequency plane \( (u,v) \), the translation vector field \( \phi_g(r,W,S) \) is isotropic (circular symmetry), i.e. the translation is the same for the same \( r = \sqrt{u^2+v^2} \). The 3D diagrams of such translation vector filed and gradient field are shown in Figure 5.
It can be seen from Figure 5 (a), for the translation vector field of equation (10), the translation at each frequency point on the concentric annulus of frequency plane \((u,v)\) is identical, and the translation \(\varphi_n(W,S,r)\) of the phase angle \(\varphi_n(u,v)\) at each frequency point \((u,v)\) in frequency domain is related to the distance \(r\) between this point and the original point only but not related to the direction (the farther the frequency point is, the larger the translation is), which is called as IPST (Isotropic Phase Stretch Transform) in this paper. The edge extracted by IPST has no direction selectivity, i.e., it is arbitrarily believed that in \(A(m,n)\), “an extreme point is the edge point (high-frequency feature point)”, which increases the translation for the compound angles of all high-frequency components and mistakes massive isolated points as the edge points, and causes that the extracted edge contains massive noises (false edge).

Therefore, an anisotropic translation vector field is proposed in this paper to make the translation to be of direction selectivity, and then the pseudo-edge is greatly reduced by the non-maximal suppression strategy to make the detection algorithm more robust to noise.

3) ANISOTROPIC TRANSLATION KERNEL FUNCTION
The typical isotropic inverse tangent translation vector field defined in equation (8) and equation (10) is irrelevant to direction because polar angle \(\theta\) is ignored in equation (6); thus, it is isotropic, i.e., it is an inverted taper surface. If the constraint of PST for “increasing A/D sampling rate to reduce time-bandwidth product of analog signal” is not considered, the PST phase warp and stretch kernel function can be designed more specifically only from the perspective of application of PST to digital signal or image processing merely.

In equation (6), if \(\theta=0,\theta=\pi/2\), the translation kernel functions of PST for direction selectivity on frequency plane \(u\) and \(v\) are obtained respectively:

\[
\begin{align*}
\varphi_1(r) &= \frac{\arctan(u)}{u} \\
\varphi_2(r) &= \frac{\arctan(v)}{v}
\end{align*}
\]

Two anisotropic (i.e., inclined to the edges of horizontal and vertical directions respectively) translation vector field prototype functions are obtained thereby:

\[
\begin{align*}
\varphi_1(u) &= u \cdot \arctan(u) - 0.5 \cdot \ln(1 + u^2) \\
\varphi_2(v) &= v \cdot \arctan(v) - 0.5 \cdot \ln(1 + v^2)
\end{align*}
\]

The phase stretch strength parameter \(S\) and warp parameter \(W\) in the non-linear warp and stretch transform are added into the two translation vector function prototypes in equation (13) to obtain the normalized phase translation vector fields \(\varphi_{1,N}(u,W,s)\) and \(\varphi_{2,N}(v,W,s)\) at each point \((u,v)\) on the frequency planes after normalization:

\[
\begin{align*}
\varphi_{1,N}(u,W,s) &= S \cdot \frac{\varphi_1(Wu)}{\varphi_1(Wu)_{\max}} \\
\varphi_{2,N}(v,W,s) &= S \cdot \frac{\varphi_2(Wv)}{\varphi_2(Wv)_{\max}}
\end{align*}
\]

3D surface and gradient surface of translation vector fields \(\varphi_{1,N}(u,W,s)\) and \(\varphi_{2,N}(v,W,s)\) are shown in Figure 6.
It can be seen from Figure 6 (a) and (c), in $\varphi_1(u)$, translation is rapidly increased with the increase of frequency component at $u$ direction on frequency plane to highlight the high frequency component at $u$ direction, thus, easily extracting the edge at horizontal direction of the image; in $\varphi_2(v)$, translation is rapidly increased with the increase of frequency component at $v$ direction on frequency plane to highlight the high frequency component at $v$ direction, thus, facilitating extraction of the edge at vertical direction of the image. The comparisons of the contour map of normalized translation vector field in equation (10) with the two in equation (14) are shown in Figure 7.

4) THEORETICAL ANALYSIS OF ANISOTROPIC KERNEL FUNCTION

Considering anisotropic kernel function $\varphi_1(u)$:

$$\varphi_1(u) = u \cdot \arctan(u) - 0.5 \cdot \ln(1 + u^2)$$  \hspace{1cm} (15)

Taylor expansion is adopted and the biquadrate of $u$ is rejected to obtain the approximate result:

$$\varphi_1(u) \approx u \cdot (u - \frac{1}{3} u^3) - \frac{1}{2} \left( u^2 - \frac{1}{2} u^4 \right) = \frac{1}{2} u^2$$  \hspace{1cm} (16)
According to the general model (5) of PST, equation (15) is adopted for PST; assuming that the source image is \( B(x,y) \) and its Fourier transformed image is \( \tilde{B}(u,v) \), without consideration of low-pass filtering, \( L(u,v)=1 \), the angle image \( A(m,n) \) is obtained:

\[
A(x,y) = \angle \text{IFFT} \left\{ \hat{B}(u,v) \cdot e^{j\phi_i(u,v)} \right\}
\]

\[
= \angle \text{IFFT} \left\{ \hat{B}(u,v) \cdot e^{j\frac{\pi}{2}} \right\}
\]

\[
= \angle \text{IFFT} \left\{ \tilde{B}(u,v) \cdot (1+j\frac{\pi}{2}) \right\}
\]

\[
= \angle \text{IFFT} \left\{ \tilde{B}(u,v) + j\frac{\pi}{2} \cdot \tilde{B}(u,v) \right\}
\]

\[
= \angle \left\{ B(x,y) + j\frac{\pi}{2} \cdot \text{IFFT} \left\{ \tilde{B}(u,v) \right\} \right\}
\]

It is noticed that:

\[
\frac{\partial^2 \tilde{B}^2}{\partial x^2} = \left( -1 \right) \cdot \left( 4\pi \right)^2 \text{IFFT} \left\{ \frac{\partial}{\partial x} \hat{B}(u,v) \right\}
\]

(18) is substituted into equation (17) to obtain:

\[
A(x,y) = \angle \left\{ \tilde{B}(u,v) - \left( \frac{j}{2} \frac{\partial^2 B(x,y)}{\partial x^2} \right) \right\}
\]

\[
= \arctan \left\{ \frac{\frac{\partial^2 B(x,y)}{\partial x^2}}{B(x,y)} \right\}
\]

\[
\approx \frac{B_x^2(x,y)}{2(4\pi)^2 B(x,y)}
\]

Equation (19) approximates the second partial derivative in x direction of some scale of source image \( B(x,y) \); local extreme point of \( A_1(x,y) \) is equivalent to the normalized/standardized local extreme point of the second-order directional partial derivative of \( B(x,y) \), which is exactly the edge point of image \( B(x,y) \) at y direction (vertical direction). The PST is written into analytical model as follows:

\[
A_1(x,y) = \text{PST}_{\phi_1} \left( B(x,y) \right) \approx \frac{B_x^2(x,y)}{B(x,y)}
\]

(20)

Similarly, it is demonstrated that:

\[
A_2(x,y) = \text{PST}_{\phi_2} \left( B(x,y) \right) \approx \frac{B_y^2(x,y)}{B(x,y)}
\]

(21)

That is, the PST transformation of the source image \( B(x,y) \) under the kernel function \( \phi_2(x,y,W,S) \) is equivalent to the second partial derivative of image \( B(x,y) \) under some normalization/scaling in the y direction; thus, the local extreme point of \( A_2(x,y) \) is exactly the edge point of image \( B(x,y) \) at x direction (horizontal direction).

It shows from equation (20) and (21) that \( A_1(x,y) \) and \( A_2(x,y) \) can be regarded as two second-order directional derivatives of image \( B(x,y) \), from which the second-order gradient map of \( B(x,y) \), \( \text{grad}A(x,y) = \sqrt{A_1^2 + A_2^2} \), is generated. Since dual edge is likely to happen during thresholding of \( \text{grad}A \), which is unfavorable for extraction of fine microorganism, the strategy shown in equation (22) is adopted.

In this paper, the strategy of “suppression of local non-maximum” was adopted for the translation kernel function with direction selectivity designed in such way, assuming that:

\[
A(m,n) = \angle \text{IFFT} \left\{ \tilde{B}(u,v) \cdot e^{j\left[ \phi_1(u,v) + \phi_2(u,v) \right]} \right\}
\]

\[
A_2(m,n) = \angle \text{IFFT} \left\{ \tilde{B}(u,v) \cdot e^{j\left[ \phi_1(u,v) + \phi_2(u,v) \right]} \right\}
\]

non-maximum suppression filtering is adopted for angle image \( A_1 \) and \( A_2 \), and the maximum value of the two is selected as the final angle image \( A(m,n) \):

\[
\text{if} (A_1(m,n) - A_1(m,n-1) > \text{threshold}) \& \&
(A_1(m,n) - A_1(m,n+1) > \text{threshold})
\]
\[\text{then } A_1(m,n) \text{ is directional edge point.}\]

\[
\text{if} (A_2(m,n) - A_2(m-1,n) > \text{threshold}) \& \&
(A_2(m,n) - A_2(m+1,n) > \text{threshold})
\]
\[\text{then } A_2(m,n) \text{ is directional edge point.}\]

\[
A(m,n) = \text{max} \{ A_1(m,n), A_2(m,n) \}
\]

(22)

Edge features are obtained through thresholding of the result image \( A(x,y) \), and the result is significantly superior to that of traditional isotropic PST, Sobel and Canny edge operator.

The experimental comparison results are shown in section IV.A.

The transform obtained by equation (22) is called as APST in this paper and can be expressed by the following mathematical model as:
APST\(B(x, y)) = \max \{ \text{PST} \varphi_1(B(x, y)), \text{PST} \varphi_2(B(x, y)) \} \tag{23} \]

D. Non-maximum suppression processing

After RTV filtering, APST was performed for the filtered image according to equation (23), in which non-maximum suppression strategy was adopted. Non-maximum suppression is an effective method to inhibit “false edge” caused by fine texture during edge feature extraction and to avoid formation of multi-pixel edge. Usually, the obtained gradient edge will have more than one pixel width. For example, the edge obtained through Sobel operator is thick and bright but the gradient map is still “vague”; while the actually required edge only has one pixel width. Non-maximum suppression is able to help reservation of maximum local gradient and inhibit all the other gradient values, which means that only the sharpest position in gradient change is reserved.

Non-maximum suppression strategy is as follows: two approximate normalized second-order gradient maps \(A_1, A_2\) are obtained after applying APST to the source image, as shown in equation (20) and (21) (it is called normalized second-order gradient map in this paper), expressed as
\[A_1(x, y) = B_1^n(x, y), A_2(x, y) = B_2^n(x, y)\]
respectively:

\[\text{If } A_1(x, y) = \max\{A_1(x-1, y), A_1(x, y), A_1(x+1, y)\}, \text{then } A_1(x, y) \text{ is edge point; otherwise } A_1(x, y) = 0;\]
\[\text{If } A_2(x, y) = \max\{A_2(x, y-1), A_2(x, y), A_2(x, y+1)\}, \text{then } A_2(x, y) \text{ is edge point; otherwise } A_2(x, y) = 0;\]

Figure 8 is the effect comparison chart with and without non-maximum suppression, and it can be seen that the edge width of the image through non-maximum suppression processing has been greatly reduced.

(a) Effect chart of Figure 2(a) without non-maximum suppression
(b) Effect chart of Figure 2(a) with non-maximum suppression
(c) Effect chart of Figure 3(a) without non-maximum suppression
(d) Effect chart of Figure 3(a) with non-maximum suppression

FIGURE 8. Effect comparison chart with and without non-maximum suppression
E. DUAL THRESHOLD SELECTION AND LAG EDGE FOLLOWING

General edge detection algorithm is to use one threshold to filter the small gradient value caused by noise or color change and to reserve large gradient value. Dual threshold, i.e., one high threshold and one low threshold were used for the algorithm in this paper to distinguish edge pixel. If the gradient value of edge pixel point is larger than the high threshold, it will be regarded as strong edge point. If gradient value of the edge pixel is smaller than the high threshold but larger than the low threshold, it will be marked as weak edge point (the effect example of strong edge point and weak edge point was shown in Figure 9). The points with gradient value smaller than the low threshold are eliminated. Strong edge point can be regarded as the true edge, while weak edge point can either be true edge or false edge caused by noise or color change. To obtain accurate results, the weak edge point caused by the latter shall be eliminated. It is generally believed that the weak edge point and strong edge point caused by true edge are connected, while it is negative to the weak edge point caused by noise. The so-called lag edge following algorithm is to detect the pixels in 8 connected domains of one weak edge point, in which once a strong edge point is found, this weak edge point will be reserved as a true edge. This algorithm will search all the connected weak edges, if any point of the connected weak edge is connected with strong edge point, this weak edge will be reserved; otherwise, it will be suppressed.

![Effect chart of strong edge point in Figure 2(a)](image1)
![Effect chart of weak edge point in Figure 2(a)](image2)
![Effect chart of strong edge point in Figure 3(a)](image3)
![Effect chart of weak edge point in Figure 3(a)](image4)

**FIGURE 9.** Effect example of strong edge point and weak edge point

III. EXPERIMENTAL RESULTS AND COMPARISON

A. COMPARISON OF EXPERIMENTAL RESULTS

Comparing analysis was performed for the edge obtained through the algorithm in this paper, i.e., image segmentation effect chart and the image segmentation by Canny, traditional isotropic PST and Sobel algorithm. The code for the algorithm of this paper was provided by the author (https://github.com/lanbaoshi110/papers), and the comparison of some images in phC-U373 data set was shown in Figure 10-14.
(a) Source image 1  
(b) Effect chart of PST processing: $S=0.16, W=1.14$  
(c) Effect chart of Canny processing

(d) Effect chart of Sobel processing  
(e) Effect chart of processing with algorithm in this paper: $S=0.16, W=1.14$  

**FIGURE 10.** Effect comparison chart 1 of all kinds of algorithms

(a) Source image 2  
(b) Effect chart of PST processing: $S=0.48, W=12.14$  
(c) Effect chart of Canny processing

(d) Effect chart of Sobel processing  
(e) Effect chart of processing with algorithm in this paper: $S=0.48, W=12.14$  

**FIGURE 11.** Effect comparison chart 2 of all kinds of algorithms
FIGURE 12. Effect comparison chart 3 of all kinds of algorithms

FIGURE 13. Effect comparison chart 4 of all kinds of algorithms
B. EXPERIMENTAL RESULTS ANALYSIS

In order to analyze performance of image segmentation algorithm, evaluation shall be carried out for all kinds of algorithms. The evaluation method for image segmentation quality is classified into two types, subjective evaluation and objective evaluation.

1) SUBJECTIVE EVALUATION

In term of subjective evaluation, segmentation quality is evaluated and judged manually. The scores of subjective evaluation given by different observers may vary greatly because every observer holds different understanding and evaluation standards for image segmentation quality; thus, it is difficult to obtain unbiased cognition of the effectiveness of segmentation algorithm. In order to reduce deviation as much as possible, it is necessary to require multiple observers to participate in the evaluation to obtain the final evaluation score for segmentation quality through comprehensive consideration of the evaluation results provided by every observer.

Subjective evaluation includes absolute subjective evaluation, relative subjective evaluation and average subjective evaluation, among which average subjective evaluation is usually used for processing of evaluation results from multiple observers [17].

Assuming that n is the evaluation level of image segmentation quality, n=1,2,…,N, and the evaluation score corresponding to the nth level of image segmentation quality is $S_n$, and $H_n$ is the number of evaluators judging that the image segmentation quality belongs to the nth level, the overall segmentation quality evaluation of the image to be evaluated is:

$$S = \sum_{n=1}^{N} S_n H_n \quad (24)$$

The total number of observers participating in the evaluation:

$$H = \sum_{n=1}^{N} H_n \quad (25)$$

The average subjective evaluation score for the segmented image was finally obtained:
Usually, there are 5 segmentation quality evaluation levels in averagely subjective evaluation, as shown in table 1:

\[
\bar{S} = \frac{\sum_{n=1}^{N} S_n \cdot H_n}{\sum_{n=1}^{N} H_n}
\]

(26)

| Segmentation quality evaluation level (n) | Segmentation effect | Score (S_n) |
|------------------------------------------|---------------------|-------------|
| Level 1                                  | Optimal             | 5 scores   |
| Level 2                                  | Favorable           | 4 scores   |
| Level 3                                  | Ordinary            | 3 scores   |
| Level 4                                  | Poor                | 2 scores   |
| Level 5                                  | Worst               | 1 score    |

The average subjective evaluation scores for image segmentation by algorithm in this paper, traditional PST, Canny and Sobel were obtained based on the number of image segmentation quality evaluators \( H = 30 \), as shown in table 2 to table 6.

**TABLE 2**

|                          | Traditional PST | Canny operator | Sobel operator | Algorithm in this paper |
|--------------------------|-----------------|----------------|----------------|-------------------------|
| \( S \)                  | 121             | 123            | 91             | 146                     |
| \( H \)                  | 30              | 30             | 30             | 30                      |
| \( \bar{S} \)            | 4.03            | 4.10           | 3.03           | 4.87                    |

**Table 3**

|                          | Traditional PST | Canny operator | Sobel operator | Algorithm in this paper |
|--------------------------|-----------------|----------------|----------------|-------------------------|
| \( S \)                  | 119             | 122            | 87             | 147                     |
| \( H \)                  | 30              | 30             | 30             | 30                      |
| \( \bar{S} \)            | 3.97            | 4.07           | 2.90           | 4.90                    |

**TABLE 4**

|                          | Traditional PST | Canny operator | Sobel operator | Algorithm in this paper |
|--------------------------|-----------------|----------------|----------------|-------------------------|
| \( S \)                  | 125             | 121            | 91             | 146                     |
| \( H \)                  | 30              | 30             | 30             | 30                      |
| \( \bar{S} \)            | 4.17            | 4.03           | 3.03           | 4.87                    |
THE AVERAGE SUBJECTIVE EVALUATION SCORES FOR FOUR ALGORITHMS INCLUDING TRADITIONAL PST ET C IN FIGURE 13

|               | Traditional PST | Canny operator | Sobel operator | Algorithm in this paper |
|---------------|-----------------|----------------|----------------|-------------------------|
| $S$           | 118             | 125            | 96             | 145                     |
| $H$           | 30              | 30             | 30             | 30                      |
| $\bar{S}$     | 3.93            | 4.17           | 3.20           | 4.83                    |

THE AVERAGE SUBJECTIVE EVALUATION SCORES FOR FOUR ALGORITHMS INCLUDING TRADITIONAL PST ET C IN FIGURE 14

|               | Traditional PST | Canny operator | Sobel operator | Algorithm in this paper |
|---------------|-----------------|----------------|----------------|-------------------------|
| $S$           | 116             | 120            | 90             | 144                     |
| $H$           | 30              | 30             | 30             | 30                      |
| $\bar{S}$     | 3.87            | 4.00           | 3.00           | 4.80                    |

2) OBJECTIVE EVALUATION

Objective evaluation can be divided into system/task-based evaluation and direct evaluation, in which direct evaluation can be further divided into analytical method and experimental method. The experimental method is also divided into the segmentation quality evaluation with supervision (reference segmentation is required) and segmentation quality evaluation without supervision (reference segmentation is not required). The classification of these evaluation methods is not mutually exclusive because each type of evaluation method has specific limitation; thus, only combination of multiple types of evaluation method can achieve mutual complementarity.

Segmentation quality evaluation with supervision is the evaluation method based on similarity or diversity measures, determining advantages and disadvantages of segmentation algorithm by measuring the similarity or diversity between algorithmic segmentation and reference segmentation. The greater the similarity or the smaller the diversity between algorithmic segmentation and reference segmentation, the higher the quality of the segmentation algorithm is. Two evaluation indexes with supervision, i.e., PRI (Probability Rand Index) [18] and GCE (Global Consistency Error) [19] were adopted in this paper for evaluation.

(1) PRI

Image segmentation can be regarded as the classification of pixel point pair either belonging to the same domain or belonging to different ones. Segmentation quality evaluation is realized by PRI through statistics of the proportion of pixels in the segmented image to be evaluated with consistent label in reference segmentation. The given segmented image to be evaluated is $S$, and $G = \{G_1, G_2, \ldots, G_k\}$ is the reference segmentation set corresponding to any pair of pixel point $(x_i, x_j)$. If their labels in $S$ and $G$ are consistent simultaneously, it means favorable segmentation effect; vice versa. The labels of $x_i$, $x_j$ in $S$ are $l_i^S$, $l_j^S$ respectively; correspondingly, their labels in $G$ are $l_i^G$, $l_j^G$ respectively. The PRI formula for $S$ and reference segmentation set $G$ is as follows:

$$PRI(S, G) = \frac{1}{C_2} \sum_{i,j(i \neq j)} \left[ I(l_i^S = l_j^G) p_{i,j} + I(l_i^S \neq l_j^G)(1 - p_{i,j}) \right]$$

(27)

Where, $N$ means total number of pixel points; $p_{i,j}$ designates the probability of the consistency of pixels $(x_i, x_j)$ labels in...
reference segmentation. In practice, it is usually set as mean value:

$$p_{ij} = \frac{1}{K} \sum_{k=1}^{K} I(i_{ij}^{(k)}, l_j^{(k)})$$  \hspace{1cm} (28)$$

According to the foregoing definition, the range of PRI is [0,1]. PRI=1 means that the segmented image to be evaluated is identical to that of reference segmentation; PRI=0 means that they are completely different. Within the range of [0,1], the higher the PRI value, the higher the segmentation quality; and vice versa.

(2) GCE

This index realizes segmentation quality evaluation based on the degree of overlapping of the region between the segmented image to be evaluated and reference segmentation image. To some extent, the segmented image can be regarded as a set of several pixel points. The evaluated segmented image is S and reference segmentation image is G. The set containing pixel $p_i$ in S is $M(S, p_i)\setminus M(G, p_i)$, and the set containing pixel $p_i$ in G is $M(G, p_i)\setminus M(S, p_i)$. Local refining error E is defined as:

$$E(S, G, p_i) = \left[ \frac{M(S, p_i)\setminus M(G, p_i)}{M(S, p_i)} \right]$$  \hspace{1cm} (29)$$

Where, operator "$\setminus$" means set difference operation. This error method is asymmetric with respect to the set of segmentation region involved in the comparison. When S is the refining region at pixel $p_i$ compared with G, $E(S, G, p_i) = 0$; when S is not intersected with G at $p_i$, $E(S, G, p_i) = 1$. According to the local refining error, GCE is defined as follows:

$$GCE(S, G) = \frac{1}{n} \min \left( \sum E(S, G, p_i) \right)$$  \hspace{1cm} (30)$$

The range of GCE is [0,1]. Smaller GCE value states higher segmentation quality; vice versa.

PRI and GCE values of image segmentation by traditional PST, Canny, Sobel and algorithm in this paper of Figure 10-14 are shown in table 7 and 8.

| TABLE 7 | PRI VALUES BY FOUR ALGORITHMS INCLUDING TRADITIONAL PST |
|---------|--------------------------------------------------------|
|         | Traditional PST | Canny | Sobel | Algorithm in this paper |
| Figure 10 | 0.81 | 0.78 | 0.72 | 0.95 |
| Figure 11 | 0.84 | 0.81 | 0.76 | 0.42 |
| Figure 12 | 0.82 | 0.84 | 0.73 | 0.96 |
| Figure 13 | 0.85 | 0.78 | 0.75 | 0.94 |
| Figure 14 | 0.82 | 0.77 | 0.73 | 0.93 |

| TABLE 8 | GCE VALUES BY FOUR ALGORITHMS INCLUDING TRADITIONAL PST |
|---------|--------------------------------------------------------|
|         | Traditional PST | Canny | Sobel | Algorithm in this paper |
| Figure 10 | 0.79 | 0.71 | 0.82 | 0.65 |
| Figure 11 | 0.78 | 0.68 | 0.81 | 0.64 |
| Figure 12 | 0.74 | 0.73 | 0.83 | 0.62 |
| Figure 13 | 0.76 | 0.74 | 0.81 | 0.65 |
| Figure 14 | 0.79 | 0.73 | 0.79 | 0.66 |

To further illustrate the advantages of the algorithm proposed in this paper, 50 AS phase-contrast images were selected from PhC-U373 data set and processed by four image segmentation algorithms including traditional PST and Canny etc. The processing results were evaluated subjectively and objectively; in the subjective evaluation, H=30 was used as the quality assessor for image segmentation, and mean value of subjective evaluation was shown in figure 15; the PRI and GCE values of objective evaluation were shown in figure 16 and 17.
It can be seen from Table 2-8 and Figure 15-17, the average subjective evaluation score and PRI value by algorithm in this paper were higher than those based on the image segmentation algorithms including traditional PST, Canny etc. GCE values are lower than those based on the image segmentation algorithms including traditional PST, Canny etc., which fully demonstrates that the noise inhibiting effect by the algorithm in this paper is better, reducing the probability of excessive, insufficient and incomplete segmentation of the activated sludge containing noise and unclear edge. In addition, its segmentation result is more approaching to the ideal edge and better segmentation effect is obtained due to enhanced image edge. The image segmentation by traditional PST showed better edge but with noise inhibiting active bacteria; Canny-based image segmentation showed the result of inaccurate segmentation effect including excessive segmentation etc., under the influence of noise; Sobel-based image segmentation showed insensitivity to image edge and error results including excessive segmentation etc.

To sum up, both the integral segmentation effect and the mean value of evaluation index further verify the effectiveness and stability of the algorithm in this paper.

IV. CONCLUSION

Aiming at the lack of direction selectivity of traditional isotropic PST, this paper put forward a kind of APST, designed the translation vector field with X and Y direction selectivity, and performed operations including non-maximum suppression and edge extraction etc., for the translation vector field at X and Y directions to obtain phase-contrast image microbial segmentation algorithm for activated sludge. The objective evaluation indexes PRI and GCE indicated that the microbial segmentation result by the algorithm proposed in this paper was significantly superior to that by traditional PST, Canny and Sobel etc. Compared with the image segmentation algorithm by traditional PST etc., the PRI and GCE evaluation indexes in the segmentation result of filamentous bacteria were promoted by around 30%; thus, the research result can be applied to digital monitoring and control of settling performance of activated sludge in wastewater treatment.

REFERENCES

[1] M.A. Molina, C.A.A. Pérez, C.A. Leiva. Characterization of Filamentous Flocs to Predict Sedimentation Parameters Using Image Analysis[J]. Journal of Sensors, 2020, 2020:1-8.
[2] Muhammad Burhan Khan, HumairaNisar, Choon Aun Ng. Generalized Modeling of the Sludge Volume Index And Suspended Solids to Monitor Activated Sludge Bioflocculation Using Image Analysis[J]. Environment Protection Engineering, 2020, Vol. 46:17-37.
[3] Jin, B, Wilen B M, Lant P, A Comprehensive Insight into Floc Characteristics and Their Impact on Compressibility and Settleability of Activated Sludge[J]. Chemical Engineering Journal, 2003 (1): 221-234.
[4] Glenn, V D, Smits K, Smets I, An Experimental Study on the Impact of Bioflocculation on Activated Sludge Separation Techniques[J]. Separation and Purification Technology, 2015(8): 94-104.
[5] Yu R F, Chen H W, Cheng W P, et al., Simultaneously Monitoring the Particle Size Distribution, Morphology and Suspended Solids Concentration in Wastewater Applying

VOLUME XX, 2020
Digital Image Analysis (DIA) [J]. Environmental Monitoring and Assessment, 2009 (4): 19-26.

[6]Jassby D, Xiao Y, Schuler A J, Biomass Density and Filament Length Synergistically Affect Activated Sludge Settling: Systematic Quantification And Modeling[J]. Water Research, 2014 (8): 457-465.

[7] Arelli A, Luccarini L, and Madoni P, Application of Image Analysis in Activated Sludge to Evaluate Correlations Between Settleability and Features of Flocs and Filamentous Species[J]. Water Science and Technology, 2009 (10): 2029-2036.

[8] Muhammad Burhan Khan, Humaira Nisar, et al. Generalized Classification Modeling of Activated Sludge Process Based on Microscopic Image Analysis[J]. Environmental Technology, 2018, Volume 39.

[9] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, P. Martinez-Gonzalez, and J. Garcia Rodriguez: Appl. Soft. Comput. 70 (2018) 41. https://doi.org/10.1016/j.asoc.2018.05.018.

[10] H. Nisar, H. Y. Hang, S. C. Siang, and M. B. Khan: Proc. 2016 IEEE Conf. Systems, Process and Control (IEEE, 2016) 102. https://doi.org/10.1109/SPC.2016.7920712.

[11] E. Shelhamer, J. Long, and T. Darrell: IEEE Trans. Pattern Anal. Mach. Intell. 39 (2017) 640. https://doi.org/10.1109/tpami.2016.2572683.

[12] O. Ronneberger, P. Fischer, and T. Brox: Proc. Medical Image Computing and Computer-Assisted Intervention, Pt II (MICCAI, 2015) 234. https://doi.org/10.1007/978-3-319-24574-4_28.

[13] Li-Jie Zhao, Shi-Da Zou, et al. Segmentation of Activated Sludge Phase Contrast Microscopy Images Using U-Net Deep Learning Model. Sensors and Materials, Vol. 31, No. 6 (2019) 2013 – 2028.

[14] Ang R B Q, Nisar H, and Khan M B, et al. Image Segmentation of Activated Sludge Phase Contrast Images Using Phase Stretch Transform[J]. Microscopy, 2018,10:1-15.

[15] Mohammad H. Asghari and Bahram Jalali, Edge Detection in Digital Images Using Dispersive Phase Stretch Transform, International Journal of Biomedical Imaging, Volume 2015, Article ID 687819, pp.1-6.

[16] Li X, Yan Q, and Xia Y, et al. Structure Extraction from Texture via Relative Total Variation[J]. ACM Trans. Graph.

Pengfei Xu was born in He Nan Province, HN, CHN in 1981. He received the M.S. in applied mathematics from Shandong University of Science and Technology, Shandong, China, in 2008. From 2008 to 2010, he was a Teaching Assistant with the Basic Teaching Department, Pingdingshan University, Henan, China. Since 2010, he has been a Lecturer with the Department of Mathematics and Information Science, Pingdingshan University, Henan province. He is the author of two books, more than five articles. His research interests include mathematical modeling, computer image processing and data analysis.

Zhiqing Zhou was born in Zhejiang province, China, in 1978. From 1998 to 2002, he received his bachelor's degree in 2002 and master's degree in 2007. Now he is working as a lecturer in School of Information Engineering of Pingdingshan University, Henan province. He has published three papers, His major fields of study include Internet of Things and Machine Vision.

Hesheng Shi is an associate professor, majoring in system simulation and experiment, IOT technology; Now he is working as a professor in School of Information Engineering of
Pingdingshan University, Henan province.

**Zexun Geng** was born in Henan Province, HN, CHN, a doctor, professor and doctoral tutor, majoring in mathematical modeling, photogrammetry and remote sensing digital image processing; unit: School of Information Engineering, Pingdingshan University; He is the author of five books, more than ten articles.