International Journal of Digital Multimedia Broadcasting

Emergent Techniques and Applications for Big Visual Data

Lead Guest Editor: Zhijun Fang
Guest Editors: Jenq-Neng Hwang, Xiaoming Huo, Hyo-Jong Lee, and Joachim Denzler

Emergent Techniques and Applications for Big Visual Data

International Journal of Digital Multimedia Broadcasting

Emergent Techniques and Applications for Big Visual Data

Lead Guest Editor: Zhijun Fang
Guest Editors: Jenq-Neng Hwang, Xiaoming Huo, Hyo-Jong Lee, and Joachim Denzler

Editorial Board

Contents

Hindawi
International Journal of Digital Multimedia Broadcasting
Volume 2017, Article ID 6468502, 2 pages
https://doi.org/10.1155/2017/6468502

Editorial

Emergent Techniques and Applications for Big Visual Data

Zhijun Fang,[1] Jenq-Neng Hwang,[2] Xiaoming Huo,[3] Hyo-Jong Lee,[4] and Joachim Denzler[5]

[1]*School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China*
[2]*Department of Electrical Engineering, University of Washington, Seattle, WA 98105, USA*
[3]*Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA*
[4]*Department of Computer Engineering, Chonbuk National University, Jeonju 54596, Republic of Korea*
[5]*Department of Mathematics and Computer Sciences, University of Jena, 07743 Jena, Germany*

Correspondence should be addressed to Zhijun Fang; zjfang@sues.edu.cn

Received 20 August 2017; Accepted 20 August 2017; Published 31 October 2017

The exponentially growing number of deployed networked cameras, from fixed to mobile and even airborne, poses strong demands on efficient techniques for better capturing, compressing, networking, computing, and visualization of these huge amounts of visual data. Emerging techniques and adherent applications of big visual data from all kinds of sensors have attracted great attention from industry, academy, and the government. Nonetheless, there are still many challenges that need to be overcome, for example, efficient techniques of analyzing big visual data and general solutions to application related difficulties. Thus, we solicited articles in this special issue that address topics covering key solutions in big visual data: sensing and capturing for big visual data; feature representation and compression of big visual data; networking and dissemination of big visual data; storage and retrieving of big visual data; security and privacy of big visual data; computing and analytics of big visual data; cloud processing and virtualization of big visual data; new applications and services of big visual data; and so forth. Before delving in the particular articles, we first share thoughts with respect to big visual data.

Data capturing and storage are the preliminary processes for big data analyses. While many researchers tried to define what big data are, the more popular description for big data arises from the properties they possess. For example, the three defining dimensions for big data are volume, variety, and velocity, which are called the Three Vs. When capturing data, the above properties should be considered in order to provide high quality data. However, digitizing and labeling a massive volume of data are time-consuming and sometimes it goes against personal privacy protection. Therefore, massive discussions and research concerning security and privacy of big visual data must be addressed as well.

A great deal of techniques are available to extract structure information from large-scale data. It is worth mentioning that machine learning is a powerful tool for identification/prediction of unknown test data by learning models from either the past history or available training data. In recent years, deep learning using convolutional neural networks (CNNs) plays a crucial role in feature representation, system modeling, and class prediction by learning from a huge amount of training data. However, it always requires expensive computations and the model has extremely high complexity which limits its universality. Recently, researchers focus on methods of reducing the computational complexity of deep learning and improving the scalability of deep learning. A typical application is face detection from large-scale real-world visual data.

Analyzing big data is in the end to improve the service for the society. New emerging applications are varying from global climate change prediction to daily life applications. For example, health care, customer behavior, intelligent transportation, recommendation system, and energy distribution. Other applications are, for example, exploring and understanding social dynamics from social media data and urban computing based on taxi trajectories to explore knowledge about a city and its citizens. In this special issue, several applications related to big visual data are investigated.

The editorial committee of this special issue received 20 submissions of which 5 papers were accepted for publication, giving an acceptance rate of 25%. Various applications and techniques related to big visual data are discussed. More specifically, Y. Yu et al. proposed a physics based 3D character animation method based on the open dynamics engine (ODE) to avoid the big data acquisition from motion capturing equipment. X. Sun et al. proposed a gradient domain image filtering-based low-light image enhancement method, which effectively extracts the illumination component of the underlying image. To efficiently deal with the problem of information overload for online users, F. Zhang et al. proposed a novel preferential diffusion movie recommendation algorithm considering the significance of the target user's nearest neighbors. Large amounts of medical CT images have been collected to establish a big medical data library. An active shape model is applied by X. Duan et al. to automatically reconstruct aortic dissection to assist operations of doctors.

Moreover, to handle large-scale text automatic processing, which uses an event network to conceal lexical relations in events and logical relations between events in document, Y. Shi et al. took advantage of the expanded Chinese Event Corpus (CEC) as data source and prior knowledge of manifestation rules of event and relation as the guide, to establish a knowledge-based rule of event manifestation, so as to achieve automatic building and improve text processing performance of event network.

As a hot research topic, more and more conferences and journals call for articles on big visual data analysis, which enhances interdisciplinary collaborations and creates new R&D domains. We believe that more and more theoretical advances and killer applications of big visual data will be available in the next years.

Zhijun Fang
Jenq-Neng Hwang
Xiaoming Huo
Hyo-Jong Lee
Joachim Denzler

Hindawi
International Journal of Digital Multimedia Broadcasting
Volume 2017, Article ID 9029315, 13 pages
https://doi.org/10.1155/2017/9029315

Research Article

Low-Light Image Enhancement Based on Guided Image Filtering in Gradient Domain

Xiankun Sun,[1,2] Huijie Liu,[2] Shiqian Wu,[3] Zhijun Fang,[2] Chengfan Li,[1] and Jingyuan Yin[1,4]

[1]*School of Computer Engineering and Science, Shanghai University, Shanghai, China*
[2]*School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai, China*
[3]*School of Machinery and Automation, Wuhan University of Science and Technology, Wuhan, China*
[4]*Earthquake Administration of Shanghai, Shanghai, China*

Correspondence should be addressed to Xiankun Sun; xksun@sues.edu.cn and Jingyuan Yin; 1061233142@qq.com

Received 2 March 2017; Revised 16 June 2017; Accepted 6 August 2017; Published 28 September 2017

Academic Editor: Fabrice Labeau

We propose a novel approach for low-light image enhancement. Based on illumination-reflection model, the guided image filter is employed to extract the illumination component of the underlying image. Afterwards, we obtain the reflection component and enhance it by nonlinear functions, sigmoid and gamma, respectively. We use the first-order edge-aware constraint in the gradient domain to achieve good edge preserving features of enhanced images and to eliminate halo artefact effectively. Moreover, the resulting images have high contrast and ample details due to the enhanced illumination and reflection component. We evaluate our method by operating on a large amount of low-light images, with comparison with other popular methods. The experimental results show that our approach outperforms the others in terms of visual perception and objective evaluation.

1. Introduction

Video surveillance is now widely used in various fields, like public security, transportation, and so on. The surveillance systems are required to perform not only in day time but also in night time. However, the videos captured in some situations such as dark place or night are very poor so that the objects can hardly be perceived by humans. Thus, it is necessary to enhance the low-light images in image processing and video surveillance.

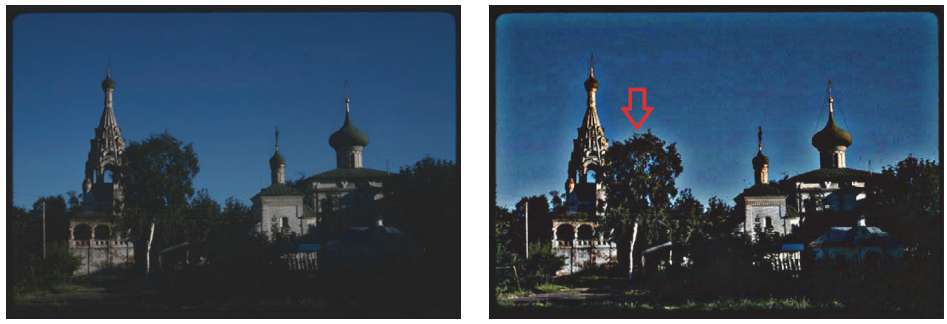
Recently, the technique of low-light image enhancement has made remarkable progress. The commonly used methods include dark channel prior model [1, 2], neural network model [3, 4], histogram equalization (HE) [5, 6], image fusion [7, 8], wavelet domain algorithm [9, 10], and illumination-reflection model [11–14]. It is noted that the adaptability of dark channel prior model is poor in disposing the images with rich details and high brightness [2]. The design and use of neural network require domain knowledge so that the underlying neural system has good generalization. The idea of histogram equalization is to merge several bins of grayscales in order to increase contrast, but such process may yield detail loss. The image fusion method needs multiple frame images, which is not applicable for single image frame. Wavelet transform is alternative technique for low-light image enhancement as shown in [9]. The most popular method in enhancing low-light images is to decompose images with the illumination-reflection model.

It was proposed by Land [15] that an image $J(x, y)$ can be decomposed by the reflection component $R(x, y)$ and the illumination component $L(x, y)$:

$$J(x, y) = R(x, y) \cdot L(x, y), \tag{1}$$

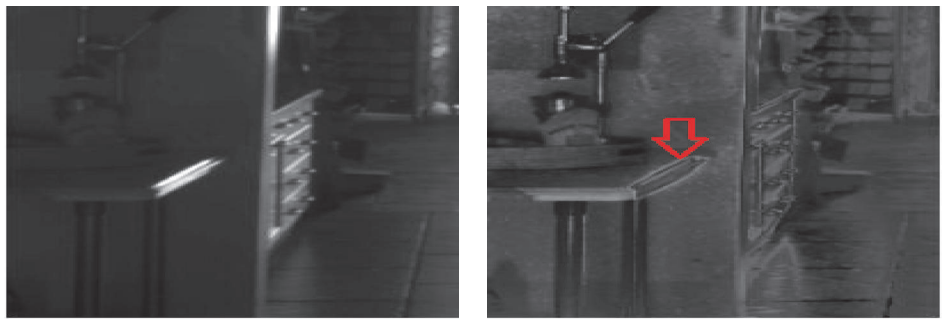
where (x, y) represents the coordinates of image. Normally, the illumination component is determined by the dynamic range of the underlying image, while the reflection component is dependent on the intrinsic characteristics of the objects within the underlying image. Equation (1) is generally converted to the logarithmic domain so that the multiplication becomes addition.

The illumination component can be obtained by various methods, such as multiscale Gaussian function in [11], an improved Gaussian function in [12], and the bilateral



(a) Input image (b) Image with halo artefact

Figure 1: Halo artefact.



(a) Input image (b) Image with gradient reversal

Figure 2: Gradient reversal.

filtering-based method in [14]. The key idea behind the Gaussian function is to use a low-pass filter as shown in (2). In the equation, $F(x, y)$ represents a center-around function, which is also a low-pass filter, and the symbol $*$ represents the convolution. Then, the reflection component can be obtained from (3) with the illumination component.

$$L(x, y) = F(x, y) * J(x, y), \tag{2}$$

$$\lg[R(x, y)] = \lg[J(x, y)] - \lg[L(x, y)]. \tag{3}$$

While calculating the illumination component and the reflection component of an image, what is most important is edge preserving and making the flat region smooth. The Multiscale Retinex (MSR) algorithm in [11] processes each color channel individually and then removes the illumination component to keep reflection component. Such processing is liable to color distortion. In addition, the method cannot achieve edge preserving and easily yields halo artefact as shown in Figure 1(b). Bilateral filtering (BLF), proposed by Tomaci and Mabduchi [16], is a good method for edge preserving and eliminating halo artefact. It was pointed out in [17] that the BLF may result in gradient reversal as shown in Figure 2(b). Accordingly, guided image filter (GIF) [17] and its variants, that is, weighted guided image filter (WGIF) [18] and guided image filter in gradient domain (GDGIF) [19] were proposed to achieve good edge preserving.

In this paper, we present a new method for low-light image enhancement, which mainly contributes to the following three aspects: (1) In order to effectively cope with the halo artefact and gradient reversal, the proposed method uses the illumination-reflection model and selects the GIF in gradient domain characterized by smoothness and edge preserving to estimate the illumination component; (2) the proposed method works on HSI color space to eliminate color distortion; (3) the illumination component and the reflection component are enhanced by nonlinear sigmoid and gamma transforms, respectively, to improve image contrast and enhance image details.

The remainder of the paper is organized as follows: Section 2 depicts the proposed low-light image enhancement algorithm. Section 3 demonstrates experimental results, followed by conclusion drawn in Section 4.

2. The Proposed Approach

The proposed method for enhancing low-light images consists of the following steps:

(1) Converting the low-light image from RGB color space to HSI color space

(2) For intensity (illumination) layer, estimating the illumination component with guided image filter in

gradient domain, followed by extracting the reflection component

(3) Enhancing the illumination component with a nonlinear sigmoid transform

(4) Enhancing the reflection component with a nonlinear gamma transform

(5) Taking the antilog of the sum of steps (3) and (4) as the enhanced intensity layer

(6) Converting the new HSI image back to RGB, which produces the final enhanced image

2.1. Guided Image Filter in Gradient Domain. GDGIF [19] and WGIF [18] are both the improvements of the guided image filter. Basically, the output image q_i is a linear transform of guided image g_i in a square window M_k centered at the pixel k:

$$q_i = a_k g_i + b_k, \quad \forall i \in M_k. \tag{4}$$

In (4), guided image g_i can be input image p_i itself. The pixel i is located in the window M_k of length $(2r + 1)$, in which r is filter radius. a_k and b_k are linear coefficients in pixel k. For GIF, WGIF, and GDGIF, the larger value of a_k implies better edge preserving. On the contrary, if the value of a_k is much closer to 0, the filters have good smoothing performance in flat regions.

According to the gradient-domain optimization framework in [20], the task of filtering an image means converting one input image p_i into the final image q_i, which can be expressed as an energy function $E(a_k, b_k)$ including the zero-order data cost function $E_d(i)$ and first-order gradient cost function $E_g(i)$ terms:

$$E(a_k, b_k) = \sum_{i \in M_k} \left(E_d(i) + E_g(i) \right), \tag{5}$$

$$E_d(i) = \left(a_k g_i + b_k - p_i \right)^2, \tag{6}$$

$$E_g(i) = \varphi(i) \left\| \nabla q_i \right\|^2, \tag{7}$$

where $\|\nabla q_i\|$ is gradient magnitude; $\varphi(i)$ is gradient weight constraint. In (4), the gradient magnitude of q_i is $\|\nabla q_i\| = a_k$.

2.1.1. Guided Image Filter. In the guided image filter [17], $\varphi_1(i) = \varepsilon$, the expressions of a_k and b_k can be obtained according to (4)–(7) shown as follows:

$$a_k = \frac{(1/|w|) \sum_{k:i \in M_k} g_i p_i - \mu_k \overline{p_k}}{\sigma_k^2 + \varepsilon}, \qquad b_k = \overline{p_k} - a_k \mu_k, \tag{8}$$

where μ_k is the mean of image g_i in the window M_k, and σ_k^2 is the variance of image g_i in the window M_k, $|w|$ is the total pixel number in window M_k, and $\overline{p_k}$ is the mean of image p_i in window M_k. ε is a regularization parameter, which controls the trade-off between edge preserving and smoothness. As ε is generally fixed ratherthan spatially varying in filtering process, halo artefact is unavoidable in edges.

2.1.2. Weighted Guided Image Filter. Li et al. [18] proposed the WGIF, where a spatially varying gradient weight constraint $\varphi_2(i) = \varepsilon / \Gamma_{g\_k}$ was added in (7).

$$\Gamma_{g\_k} = \frac{1}{N} \sum_{i=1}^{N} \frac{\sigma_{g,1}^2(k) + \lambda}{\sigma_{g,1}^2(i) + \lambda}, \tag{9}$$

where Γ_{g_k} is a single-scale edge-aware weighting, which is defined by using local variances in 3×3 windows, $\sigma_{g,1}^2(k)$ is the variance of guided image g_i in the 3×3 window, N is the number of image pixels, λ is a small constant and its value is $(0.001 \times L)^2$, and L is the dynamic range of the input image. The expressions of a_k' and b_k' in the WGIF can be obtained according to (4)–(7) as shown below:

$$a_k' = \frac{(1/|w|) \sum_{k:i \in M_k} g_i p_i - \mu_k \overline{p_k}}{\sigma_k^2 + \varepsilon / \Gamma_{g\_k}}, \qquad b_k' = \overline{p_k} - a_k' \mu_k. \tag{10}$$

It is noted that the edge-aware weighting Γ_{g_k} is spatially varying in the WGIF; that is, Γ_{g_k} is larger than 1 when the pixel k locates in the edge area, and Γ_{g_k} is smaller than 1 when the pixel k locates in the flat area. As a result, a_k' is closer to 1 than a_k in the GIF, which implies that WGIF has the better edge preservation than GIF.

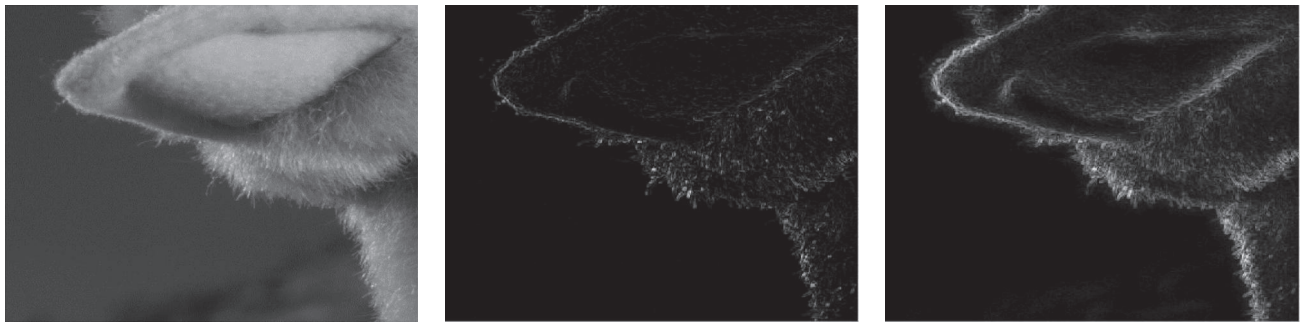
WGIF can reduce the halo artefact to some extent. However, both the GIF and the WGIF filters have no explicit constraints to cope with edges. Both cannot preserve edges well because image filtering is performed on edges, which definitely smoothed the edges [19, 21].

2.1.3. Guided Image Filter in Gradient Domain. Kou et al. [19] proposed the guided image filter in gradient domain (GDGIF), by adding an explicit first-order (gradient domain) edge-aware constraint $E_e(i)$ to gradient-domain equation (5) shown below:

$$E(a_k, b_k) = \sum_{i \in M_k} \left(E_d(i) + E_g(i) + \phi E_e(i) \right), \qquad E_e(i) = \alpha \left\| \nabla q_i - \nabla p_i \right\|^2, \tag{11}$$

where $E_e(i)$ is edge-aware constraint. The aim is to perceive the changes in local neighbourhoods so that the similar filtering is performed in the similar regions. α is a weight value, ∇q_i and ∇p_i represent the gradient values of output image q_i and input image p_i.

In the GDGIF, a multiscale edge-aware varying spatially gradient weight constraint $\varphi_3(i) = \varepsilon / \widehat{\Gamma}_{g\_k}$ is added in (7). The energy function of GDGIF is shown in (12), in which the



(a) Original image (b) Edge results by Γ_{g_k} (c) Edge results by $\widehat{\Gamma}_{g\_k}$

Figure 3: Comparison of $\widehat{\Gamma}_{g\_k}$ and Γ_{g_k}.

second item is the combination of the first-order gradient cost function $E_g(i)$ and the edge-aware constraint $E_e(i)$.

$$E\left(a_k'', b_k''\right) = \sum\left(\left(a_k'' g_i + b_k'' - p_i\right)^2 + \frac{\varepsilon}{\widehat{\Gamma}_{g\_k}}\left(a_k'' - \gamma_k\right)^2\right), \tag{12}$$

$$\widehat{\Gamma}_{g\_k} = \frac{1}{N}\sum_{i=1}^{N}\frac{\chi_k + \lambda}{\chi_i + \lambda}, \tag{13}$$

$$\gamma_k = 1 - \frac{1}{1 + e^{\eta(\chi_k - \mu_{\chi,\infty})}}, \tag{14}$$

where $\widehat{\Gamma}_{g\_k}$ is defined by local variances of $(2r+1)\times(2r+1)$ windows of all pixels; r is the filter radius. It is noted that $\widehat{\Gamma}_{g\_k}$ is a multiscale edge-aware weighting varying spatially, in which $\chi_k = \sigma_{g,1}(k) * \sigma_{g,r}(k)$. The edge-aware weighting detects the edge more accurately, and a pixel is detected as an edge pixel when the two scale variances are very large.

The comparison of $\widehat{\Gamma}_{g\_k}$ and Γ_{g_k} in WGIF of an image is shown in Figure 3. The edges of the image are detected accurately by using the multiscale edge-aware weighting $\widehat{\Gamma}_{g\_k}$ rather than using the single-scale edge-aware weighting Γ_{g_k} in WGIF.

In (14), γ_k is an edge-aware constraint to preserve edges. $\mu_{\chi,\infty}$ is the mean of all χ_i, and the value of η is $4/(\mu_{\chi,\infty} - \min(\chi_i))$. The value of γ_k is close to 1 when the pixel k locates in the edge area, and the value is close to 0 when the pixel k locates in the flat area.

The expressions of a_k'' and b_k'' in the GDGIF can be obtained according to (12)–(14), shown as follows:

$$a_k'' = \frac{(1/|w|)\sum_{k:i\in M_k} g_i p_i - \mu_k \overline{p_k} + \left(\varepsilon/\widehat{\Gamma}_{g\_k}\right)\gamma_k}{\sigma_k^2 + \varepsilon/\widehat{\Gamma}_{g\_k}}, \qquad b_k'' = \overline{p_k} - a_k'' \mu_k. \tag{15}$$

When the input image p_i and guided image g_i are the same, the GDGIF has better edge preserving and smoother features than the GIF and the WGIF due to the following two points.

(1) When pixel k locates in the edge area, the expression of a_k'' can be computed as

$$a_k'' = \frac{\sigma_k^2 + \left(\varepsilon/\widehat{\Gamma}_{g\_k}\right)\gamma_k}{\sigma_k^2 + \varepsilon/\widehat{\Gamma}_{g\_k}}. \tag{16}$$

It is seen that the value of γ_k is 1, and the value of a_k'' is 1, which is independent of parameter ε. In fact, the value of a_k'' in GDGIF is much closer to 1 than it is in GIF and WGIF. Hence, GDGIF has the best edge preserving feature.

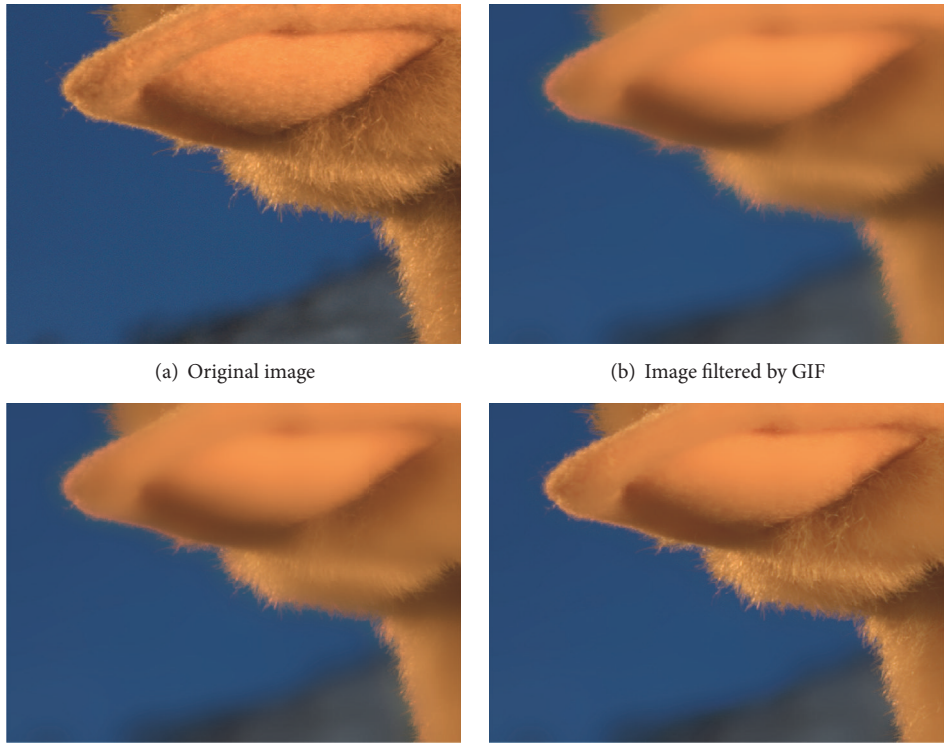
(2) When a pixel k locates in the flat areas, the value γ_k is close to 0, the parameter a_k'' is accordingly independent of the choice of ε. As a result, we can select larger ε in GDGIF than that in WGIF and GIF, so that better smoothing is achieved without affecting the edge preserving [19].

For example, we use GIF, WGIF, and GDGIF for image filtering by choosing the same filter radius $r = 16$ and regularization parameter $\varepsilon = 0.8\text{^}2$. The result is shown in Figure 4. It is observed that the image filtered by GDGIF has the best edge preserving.

In summary, the GDGIF has the best edge preserving and smoothing features. As a result, we choose the GDGIF to estimate the illumination component.

2.2. Enhancing the Intensity Layer. Generally, the processing of an RGB image is to operate the R, G, and B three channels separately, which is time-consuming. In this work, the low-light image is converted from RGB color space to HSI color space. The HSI color space stems from the human visual system, using three elements of color, hue (H), saturation (S), and intensity (I) to describe the color, which is more consistent with human visual feature than RGB color space. In HSI color space, the intensity component is enhanced while the hue and saturation components are extracted without further processing.

2.2.1. Estimating Illumination Component. Based on illumination-reflection model, combined with (2), (4), and (15), we use the GDGIF to estimate the illumination



(a) Original image (b) Image filtered by GIF

(c) Image filtered by WGIF (d) Image filtered by GDGIF

Figure 4: The difference between GIF, WGIF, and GDGIF.

component $L_I(x, y)$ of intensity layer image $I(x, y)$ as follows:

$$L_I(x, y) = q(x, y) * I(x, y), \tag{17}$$

where $q(x, y)$ is represented by the GDGIF function, as shown in (4). Then the reflection component $R_I(x, y)$ of intensity layer image can be obtained according to (3) and (17) as follows:

$$\begin{aligned} r_I(x, y) &= \lg[R_I(x, y)] \\ &= \lg[I(x, y)] - \lg[L_I(x, y)]. \end{aligned} \tag{18}$$

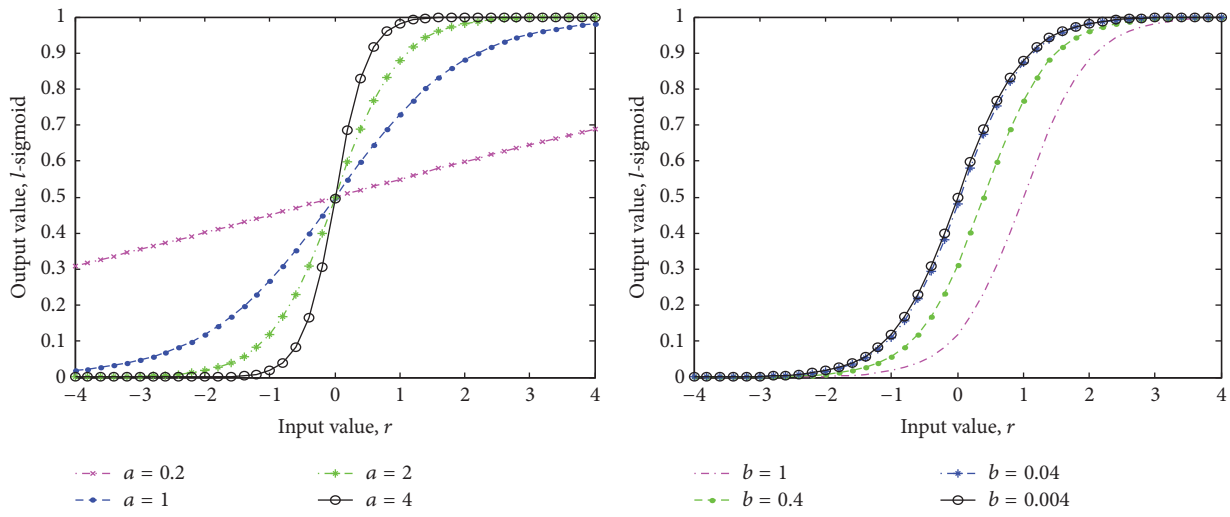
2.2.2. Enhancing Illumination Component. Normally, the methods based on illumination-reflection model are to extract and then enhance the reflection component without considering the illumination component. Such processing leads to a lack of coordination between gray levels and yields color distortion. To cope with this problem, Wu et al. [13] proposed enhancing the reflection component together with the illumination component. Inspired by this idea, we process the illumination component by the nonlinear tensile sigmoid transform as shown in Figure 5. It is indicated in [22] that the sigmoid transform has the ability to sharpen images, highlight the local details, and stretch the image contrast. The nonlinear tensile sigmoid transform is expressed in the following equations:

$$\begin{aligned} l_{\text{Sigmoid}}(x, y) &= r_{\min} + \frac{r_{\max} - r_{\min}}{1 + e^{-a(r(x,y) - b\cdot(r_{\max} - r_{\min}))}}, \\ l_{I\_\text{out}}(x, y) &= l_I(x, y) \times l_{\text{Sigmoid}}(x, y), \end{aligned} \tag{19}$$

where l_{Sigmoid} is self-defined sigmoid nonlinear function, whose range is $[0, 1]$, $r_{\min}$ and $r_{\max}$ represent the minimum and maximum values of reflection component, respectively, and $l_I(x, y)$ and $l_{I\_\text{out}}(x, y)$ are the initial and the enhanced illumination component, respectively. It is noted that there are two important parameters a, b in which the parameter a controls how an image is enhanced and the parameter b controls the contrast enhancement. Generally, a large value of a enhances an image greatly. On the other hand, a small value of b enhances the contrast of the dark regions, and a large value of b enhances the contrast of the light regions. Figure 6 shows examples by taking different parameters. In this study, the parameters a, b are selected as 2 and 0.004, respectively.

2.2.3. Enhancing the Reflection Component. It is known that human eyes are not sensible for high brightness difference but sensible for the small difference in low intensity. Thus, gamma transform [23] is normally employed to enhance the reflection component as follows:

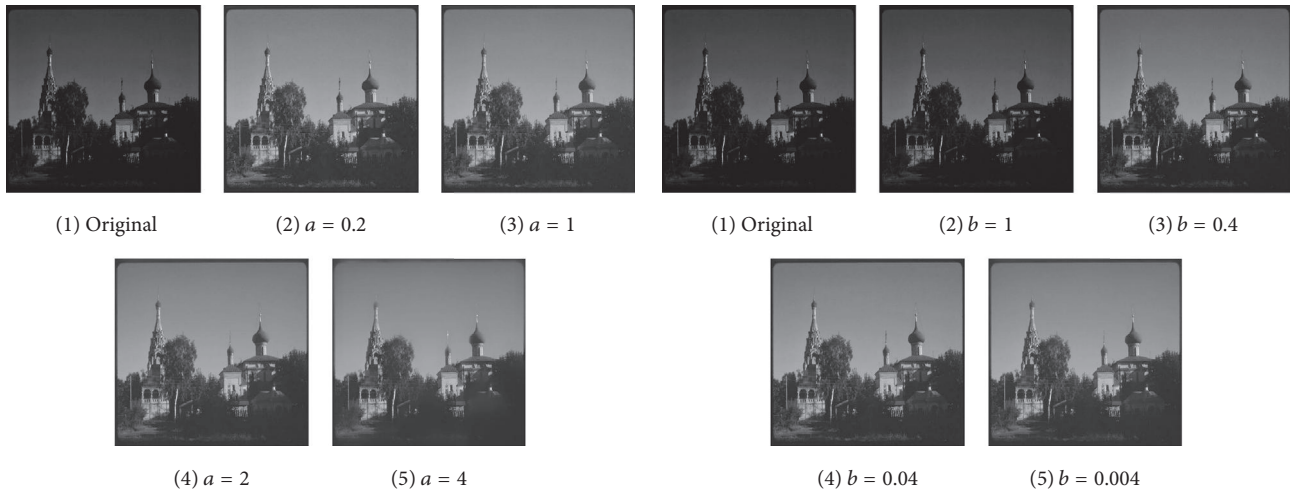
$$r_{I\_\text{out}}(x, y) = c r_I^{\gamma}(x, y), \tag{20}$$



(a) Sigmoid curves in different a values ($b = 0.004$)

(b) Sigmoid curves in different b values ($a = 2$)

Figure 5: Sigmoid function curves.



(a) Illumination enhancement in different a values ($b = 0.004$)

(b) Illumination enhancement in different b values ($a = 2$)

Figure 6: Illumination component enhanced by sigmoid curves.

where c is a positive constant; γ is a parameter to control the image contrast.

It is seen from Figure 7 that when γ is less than 1, the contrast in low intensities is increased. On the contrary, the contrast in high intensities is enhanced in case $\gamma > 1$. The effect of parameter c is shown in Figure 8. In this work, the parameters are experimentally selected as follows: $\gamma = 0.6$; $c = 0.3$.

2.3. Final Enhanced Image. By taking antilog of the illumination component combined with the reflection component, we obtain the enhanced intensity image. The new HSI image is made up of enhanced intensity layer, original hue, and saturation layer. Then the enhanced HSI image is converted into RGB image to obtain the final enhanced image. The whole process is shown in Figure 9.

3. Experimental Results and Discussions

As there is no public low-light image database, we collect 20 images from the Library of Congress and Internet as shown in Figure 10. In the simulation, the parameters are predefined as follows: window radius $r = 16$, regularization parameter $\varepsilon = 0.001$, the sigmoid parameters $a = 2$, $b = 0.004$, and gamma parameters $\gamma = 0.6$, $c = 0.3$. All experiments are performed in Matlab code on Windows 7 operation system. The computer is Intel® core™ i5-4570 3.20 GHz, and the RAM is 4.00 GB. The popular algorithms such as traditional

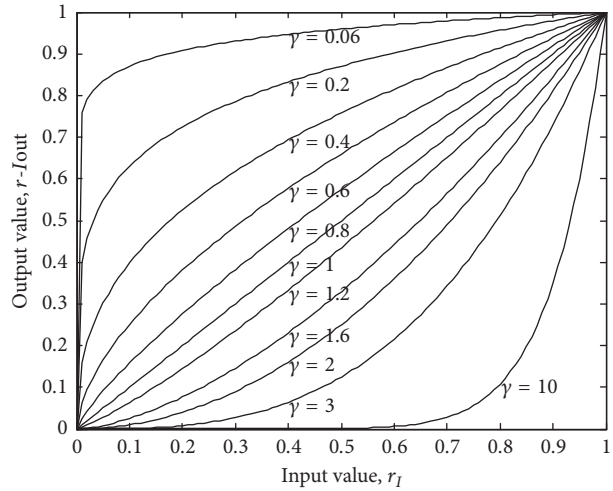


Figure 7: Gamma transforms ($c = 1$).

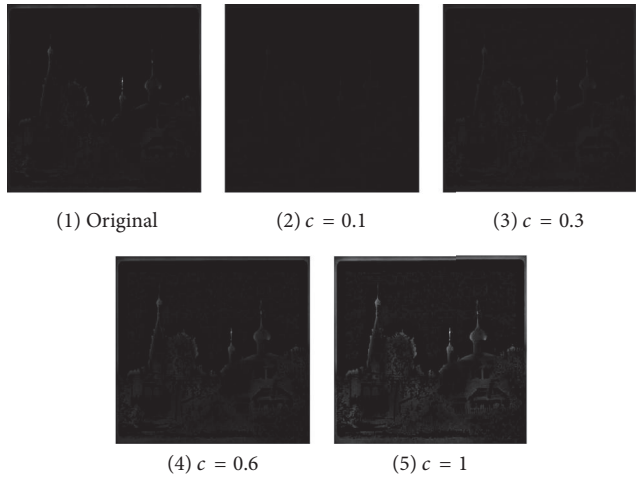


Figure 8: Reflection component enhanced by different c values.

MSR [11], Hao's algorithm [14], He's algorithm [17], histogram equalization (HE), improved MSR [12], and Kim's algorithm [9] are implemented for performance comparison.

3.1. Subjective Evaluation. It is observed from Figures 11 and 12 that the images enhanced by traditional MSR algorithm have obvious "white" phenomenon, which indicates color distortion. The images enhanced by Hao's method result in gradient reversal artefact as seen in red square from image 4 (white tower) in Figure 11, which indicates that the bilateral filtering has poor edge preserving property. Moreover, the resulting images by Hao's method are blurred which can be seen from images 1 (church) and 5 (the study) in Figure 11. He's algorithm also shows halo artefact which can be seen from image 1 (church), and this method cannot enhance the image contrast and brightness shown in Figure 12.

Histogram equalization algorithm can enhance image contrast, but this method misses details and enlarges noise, which can be observed from Figures 11 and 12. Improved MSR algorithm is better than the traditional MSR algorithm, but this method causes halo artefact which can be seen from image 1 (church) in Figure 11 (highlighted box). Moreover, "white" phenomenon exists as shown in image 13 (boy), where the tire is over-enhanced and loses detailed information. Our results show that improved MSR algorithm is not effective in dealing with night images, as observed in Figure 12.

Kim's algorithm also causes color distortion, which can be seen from image 1 (church) in Figure 11, and the sky in the red square becomes gray. The enhanced images may be blurred as seen in image 1 (church), which implies that Kim's algorithm has poor edge preserving.

As seen from Figures 11 and 12, the proposed algorithm has the best color fidelity and edge preserving. Also, the enhanced images have less noise and clearer details than the one enhanced by histogram equalization algorithm.

3.2. Objective Evaluation. Currently, no standard objective metrics are proposed for enhancement assessment of low-light images. In this study, we employ information entropy

(a) Original (b) Intensity layer image

(c) Illumination component (d) Reflection component

(e) Enhanced illumination component (f) Enhanced reflection component

(g) Enhanced intensity image (h) Enhanced RGB image

FIGURE 9: The proposed procedure of low-light images.

(IE) to evaluate image details and choose the average edge intensity (AEI) to evaluate the edge preserving feature of enhanced images. The IE is defined as follows:

$$\mathrm{IE} = -\sum_{s=0}^{L-1} P(r_s) \log_2 P(r_s), \tag{21}$$

where $P(r_s)$ represents the probability of gray level s; L is the total number of gray levels. Larger IE indicates more information, which implies that there are more details in the underlying image.

The AEI is defined as follows:

$$\mathrm{AEI} = E(g(i,j)) = E\left(\sqrt{S_x(i,j)^2 + S_y(i,j)^2}\right), \tag{22}$$

where $S_x(i,j)$ and $S_y(i,j)$ represent gradients in x and y directions on image edges. E represents expectation. In this study, Sobel operator is used for gradient computation. Larger

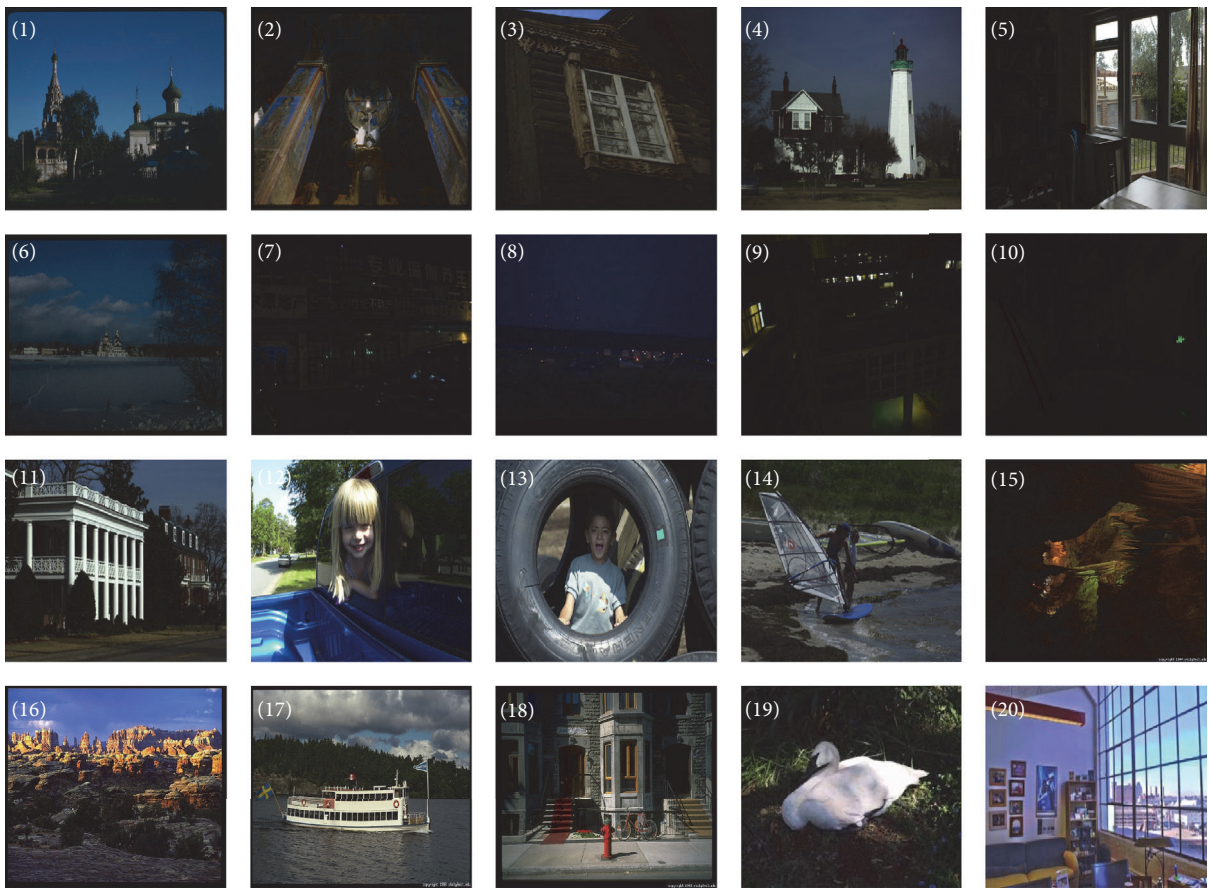


Figure 10: Datasets used in experiments.

AEI indicates good edge preserving. Also, the efficiency of the proposed method is compared with other methods.

Table 1 shows the average IE and AEI and operating time based on 20 test images. It is observed that the performance of He's method is not good in terms of IE. The images enhanced by He's method are usually dark and lose many details. Figure 13 shows the IE performances of the best three methods, that is, Hao's algorithm, Kim's algorithm, and the proposed method. It is seen that the median IE (marked by a red line) of the proposed method is the largest. The mean IE (marked by a green line) of the proposed method is 9.8% higher than Hao's algorithm and 1.8% higher than Kim's algorithm in the 20 test images.

On the other hand, it is observed from Table 1 that the improved MSR, histogram equalization (HE), Kim's algorithm, and the proposed algorithm achieve top four performances in terms of AEI. It is noted that high AEI in the histogram equalization accounts for much noise. The AEI of the improved MSR, Kim's method, and the proposed method on each individual image are shown in Figure 14. Results show that the median AEI (marked by a red line) of the proposed method is the largest, and the mean AEI (marked by a green line) of the proposed method is 26.8% higher than the improved MSR and 15.6% higher than Kim's algorithm.

Table 1: The average IE, AEI, and operating time (sec) on twenty images.

	Average IE	Average AEI	Average time
Original	11.057	50.756	N/R
Traditional MSR	13.439	67.485	1.77
Hao et al.'s [14]	14.091	57.838	16.289
He et al.'s [17]	11.867	61.952	0.377
HE	14.761	83.874	0.281
Improved MSR [12]	13.234	83.606	7.915
Kim et al.'s [9]	15.193	96.477	0.867
The proposed method	15.478	114.297	0.698

In summary, the proposed method achieves the best performance in terms of both IE and AEI. Furthermore,

Figure 11: Comparison among different algorithms.

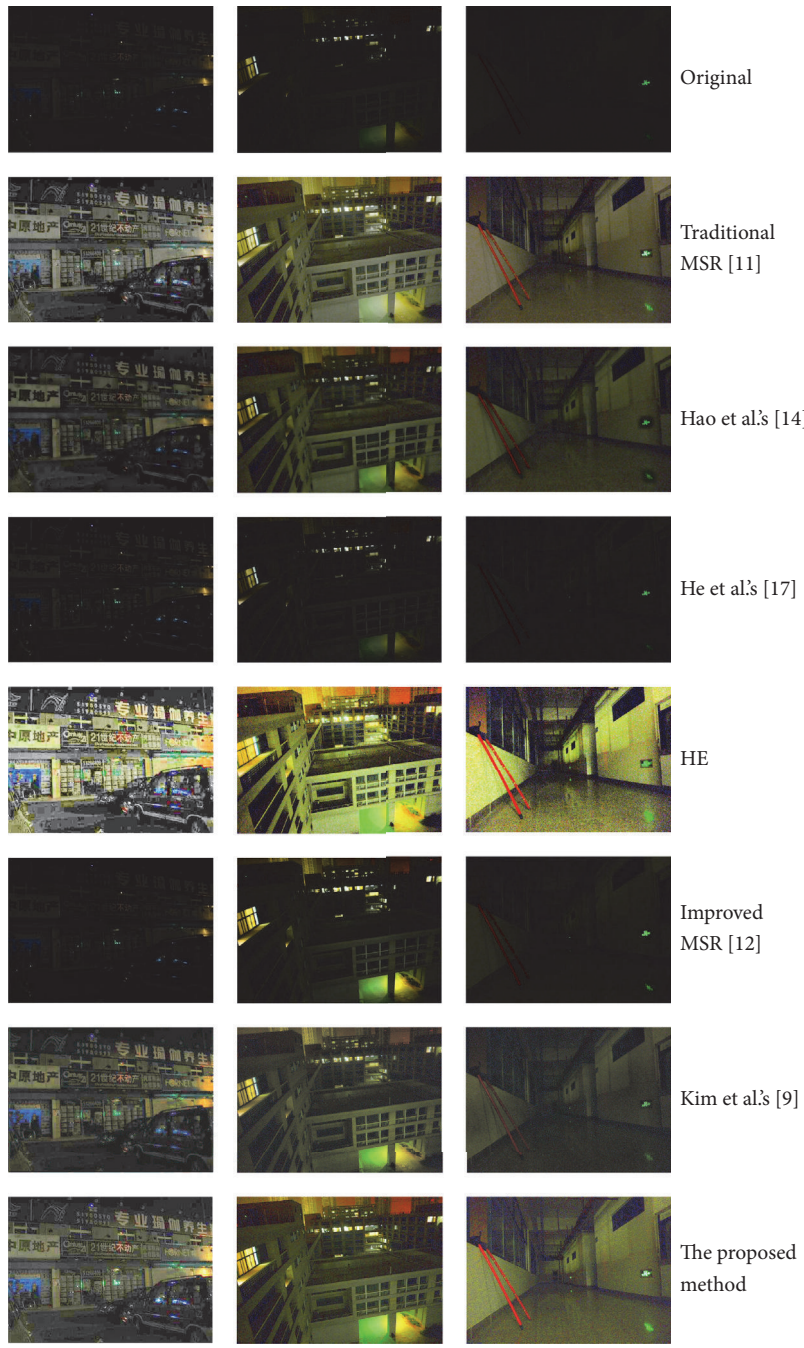


Figure 12: Results by different algorithms.

the proposed method is very efficient and achieves the 3rd position among the 7 methods as shown in Table 1. It is noted that He's algorithm has the worst enhanced image quality. Figure 15 shows the median operating time (marked by a red line) and the mean operating time (marked by a green line) of HE, Kim's, and the proposed method. Generally, the bilateral filtering is not as efficient as the GDGIF. The computation complexity of bilateral filtering is $o(Nr^2)$, where r refers to the window radius of bilateral filtering, and N is the number of image pixels while the GDGIF's computation complexity is $o(N)$. In other words, the GDGIF's computation complexity is not related to the filter size. Results show that the average time by the proposed method is 59.7% higher than HE and 19.4% less than Kim's algorithm.

4. Conclusion

A low-light image enhancement algorithm is presented in the paper. By decomposing a low-light image into the illumination component and the reflection component, it

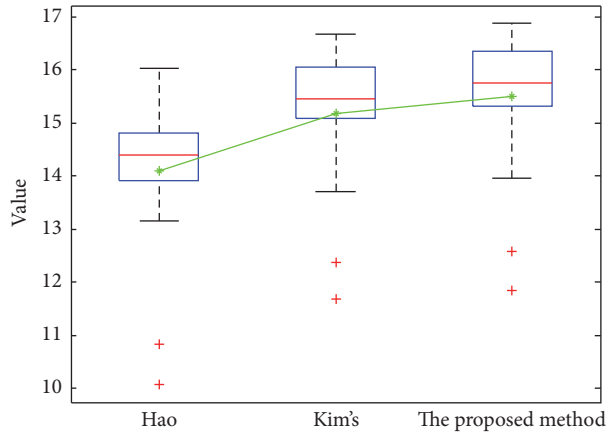


Figure 13: IE results by Hao's, Kim's, and the proposed method.

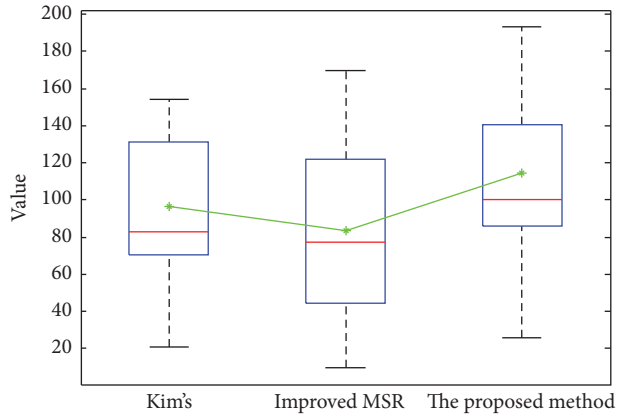


Figure 14: AEI results by Kim's, improved MSR, and the proposed method.

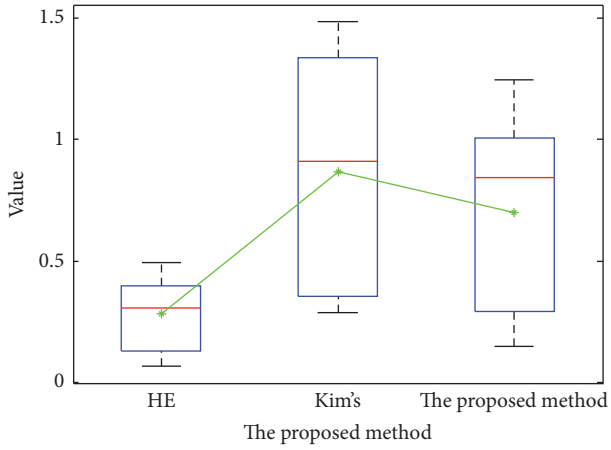


Figure 15: Operating time by HE, Kim's, and the proposed method.

offers a solution to expand illumination and enhances image details separately. Specifically, the illumination component is processed using guided image filter in gradient domain, followed by nonlinear sigmoid transform. The reflection component is enhanced by the gamma transform. This solution enhances low-light images and effectively avoids distortions (for example color) and annoying artefacts (e.g., blurring, halo). Then, the final result is obtained by antilogging the sum of the enhanced two components. Experimental results demonstrate that the enhanced images by the proposed method are visually pleasing by subjective test and the performance of the proposed method outperforms the existing methods in terms of both IE and AEI assessments. Moreover, the proposed algorithm is efficient because the computation complexity is not related to filter size. The proposed method has great potential to implement in real-time low-light video processing.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61371190, Natural Science Foundation of Shanghai under Grants 13ZR1455200 and 17ZR1411900, the Opening Project of Shanghai Key Laboratory of Integrated Administration Technologies for Information Security (AGK2015006), the Founding Program for the Cultivation of Young University Teachers of Shanghai (ZZGCD15090), and the Research Start-Up Founding Program of Shanghai University of Engineering Science (2016-56). The authors would like to thank Se Eun Kim for sharing his code.

References

[1] A. Alajarmeh, R. A. Salam, M. F. Marhusin, and K. Abdulrahim, "Real-time video enhancement for various weather conditions using dark channel and fuzzy logic," in *Proceedings of the 2014 International Conference on Computer and Information Sciences, ICCOINS 2014*, mys, June 2014.

[2] X. Jiang, H. Yao, S. Zhang, X. Lu, and W. Zeng, "Night video enhancement using improved dark channel prior," in *Proceedings of the 2013 20th IEEE International Conference on Image Processing, ICIP 2013*, pp. 553–557, aus, September 2013.

[3] L. Tang, S. Chen, W. Liu, and Y. Li, "Improved Retinex image enhancement algorithm," in *Proceedings of the 2011 2nd International Conference on Challenges in Environmental Science and Computer Engineering, CESCE 2011*, pp. 208–212, chn, December 2011.

[4] M. Athimethphat and K. Kritayakirana, "Enhanced illumination balancing with neural network for improved degraded scanned text-photo images," in *Proceedings of the 8th Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI) Association of Thailand - Conference 2011, ECTI-CON 2011*, pp. 983–986, tha, May 2011.

[5] G. Yadav, S. Maheshwari, and A. Agarwal, "Contrast limited adaptive histogram equalization based enhancement for real time video system," in *Proceedings of the 3rd International Conference on Advances in Computing, Communications and Informatics, ICACCI 2014*, pp. 2392–2397, ind, September 2014.

[6] Z. Rong, Z. Li, and L. Dong-Nan, "Study of color heritage image enhancement algorithms based on histogram equalization," *Optik*, vol. 126, no. 24, pp. 5665–5667, 2015.

[7] J. Xiao, H. Peng, Y. Zhang, C. Tu, and Q. Li, "Fast image enhancement based on color space fusion," *Color Research and Application*, vol. 41, no. 1, pp. 22–31, 2016.

[8] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiére, and W. Wu, "Objective assessment of multiresolution image fusion algorithms for context enhancement in Night vision: A comparative study," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 94–109, 2012.

[9] S. E. Kim, J. J. Jeon, and I. K. Eom, "Image contrast enhancement using entropy scaling in wavelet domain," *Signal Processing*, vol. 127, pp. 1–11, 2016.

[10] E. Provenzi and V. Caselles, "A wavelet perspective on variational perceptually-inspired color enhancement," *International Journal of Computer Vision*, vol. 106, no. 2, pp. 153–171, 2014.

[11] D. J. Jobson, Z.-U. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 965–976, 1997.

[12] A. B. Petro, C. Sbert, and J. Morel, "Multiscale Retinex," *Image Processing On Line*, vol. 4, pp. 71–88, 2014.

[13] S. Wu, Z. Hu, W. Yu, and J. Feng, "An improved image enhancement approach based on Retinex theory," in *Proceedings of the 2013 International Conference on Information Technology and Applications, ITA 2013*, pp. 67–71, chn, November 2013.

[14] W. Hao, M. He, H. Ge, C.-J. Wang, and Q.-W. Gao, "Retinex-like method for image enhancement in poor visibility conditions," in *Proceedings of the 2011 International Conference on Advanced in Control Engineering and Information Science, CEIS 2011*, pp. 2798–2803, chn, August 2011.

[15] E. Land, "The retinex," *American Scientists*, vol. 52, no. 2, pp. 247–264, 1964.

[16] C. Tomaci and R. Mabduchi, "Bilateral filtering for gray and color images," in *Proceedings of the Proceeding of the 6th IEEE International Conference Computer Vision*, pp. 839–846, Bombay, India, January 1998.

[17] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.

[18] Z. Li, J. Zheng, Z. Zhu, W. Yao, and S. Wu, "Weighted guided image filtering," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 120–129, 2015.

[19] F. Kou, W. Chen, C. Wen, and Z. Li, "Gradient domain guided image filtering," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4528–4539, 2015.

[20] P. Bhat, C. L. Zitnick, M. Cohen, and B. Curless, "GradientShop: A gradient-domain optimization framework for image and video filtering," *ACM Transactions on Graphics*, vol. 29, no. 2, article no. 10, 2010.

[21] M. Hua, X. Bie, M. Zhang, and W. Wang, "Edge-aware gradient domain optimization framework for image filtering by local propagation," in *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014*, pp. 2838–2845, usa, June 2014.

[22] O. Nuri and C. Ender, "A non-linear technique for the enhancement of extremely non-uniform lighting images," *Journal of Aeronautics and Space Technologies*, vol. 3, no. 3, pp. 37–47, 2007.

[23] A. K. Bhandari, A. Kumar, and G. K. Singh, "Improved knee transfer function and gamma correction based method for contrast and brightness enhancement of satellite image," *AEU - International Journal of Electronics and Communications*, vol. 69, no. 2, pp. 579–589, 2015.

Hindawi
International Journal of Digital Multimedia Broadcasting
Volume 2017, Article ID 3163759, 8 pages
https://doi.org/10.1155/2017/3163759

Research Article

Visual Three-Dimensional Reconstruction of Aortic Dissection Based on Medical CT Images

Xiaojie Duan,[1] Dandan Chen,[1] Jianming Wang,[1] Meichen Shi,[1] Qingliang Chen,[2] He Zhao,[1] Ruixue Zuo,[1] Xiuyan Li,[1] and Qi Wang[1]

[1]*Tianjin Key Laboratory of Optoelectronic Detection Technology and Systems, School of Electronics and Information Engineering, Tianjin Polytechnic University, Tianjin 300387, China*
[2]*Tianjin Chest Hospital, Tianjin 300000, China*

Correspondence should be addressed to Jianming Wang; wangjianming@tjpu.edu.cn

Received 27 February 2017; Revised 7 May 2017; Accepted 7 June 2017; Published 19 July 2017

Academic Editor: Zhijun Fang

With the rapid development of CT technology, especially the higher resolution of CT machine and a sharp increase in the amount of slices, to extract and three-dimensionally display aortic dissection from the huge medical image data became a challenging task. In this paper, active shape model combined with spatial continuity was adopted to realize automatic reconstruction of aortic dissection. First, we marked aortic feature points from big data sample library and registered training samples to build a statistical model. Meanwhile, gray vectors were sampled by utilizing square matrix, which set the landmarks as the center. Posture parameters of the initial shape were automatically adjusted by the method of spatial continuity between CT sequences. The contrast experiment proved that the proposed algorithm could realize accurate aorta segmentation without selecting the interested region, and it had higher accuracy than GVF snake algorithm (93.29% versus 87.54% on aortic arch, 94.30% versus 89.25% on descending aorta). Aortic dissection membrane was extracted via Hessian matrix and Bayesian theory. Finally, the three-dimensional visualization of the aortic dissection was completed by volume rendering based on the ray casting method to assist the doctors in clinical diagnosis, which contributed to improving the success rate of the operations.

1. Introduction

Aortic dissection (AD) is a cardiovascular disease that is a dangerous threat to human health, which can quickly lead to death [1]. The main reason for this disease is that tissue weakness and high blood pressure lead to one or more aortic tissues perforation(s), blood flow along the intimae resulting in two separate blood flow channels: the true lumen (the primary aorta bed of blood flow) and the false lumen (a channel entirely within the media which appears during an aortic dissection) [2], as shown in Figure 1.

At present, the main separation therapy of aortic dissection is the lumen isolation technique requiring that the clinicians can clearly know the crevasse position, range, quantity, severity, and so on before surgery. In order to improve positive rate of aortic dissection, realize automation guidance to interventional treatment, and achieve precise surgery or postoperative evaluation, the aortic dissection 3D reconstruction system is indispensable. The key technology is to achieve aortic segmentation. Threshold-based methods of image segmentation are challenged by intensity gradients within the image volume [3].

And edge detection methods are challenged by poor contrast in the medical images. Methods based on specific theory introduce mathematical models into image segmentation areas, and active shape model (ASM) belongs to the strategy of "top-down"; it combines the prior knowledge of top floor with information at the bottom of the image characteristics and is able to achieve accurate segmentation of complex medical images [4]. In three-dimensional space, there is continuity between CT image sequences; the thicker the slice is, the robuster the continuity is kept between CT slices. Based on the above methods, this paper adopts the method that combines active shape model with spatial continuity to extract aorta area quickly and accurately, which eliminates

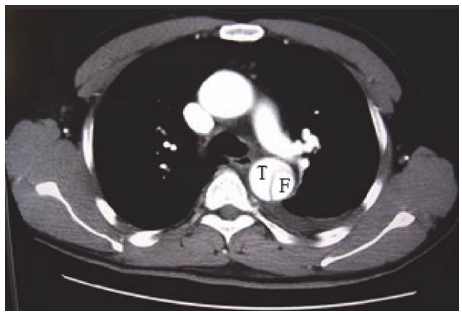


Figure 1: The true and the false lumen.

the interference of other organs by shape constraints in the process of extracting the aorta. Then extract aortic dissection membrane using Hessian matrix. And finally, we make use of volume rendering based on ray casting algorithm to perform three-dimensional reconstruction of the aortic dissection. Moreover, we set the transparency and colors of aortic dissection 3D model to make aortic dissection more intuitive; the location and the size of the intimae crevasse can be easily obtained.

2. Principle of System

The principle framework of the proposed three-dimensional reconstruction system is shown in Figure 2, via segmentation, extraction, and reconstruction of aortic dissection patients' CT images, to realize three-dimensional visualization of aortic dissection. Firstly, we preprocess the original CT images in order to eliminate noise and adjust the images' brightness for subsequent processing. Secondly, the active shape model is used to segment the aorta region, combined with the spatial continuity of the huge number of CT image sequences, automatically adjusting the posture parameters of the initial shape of each layer to improve the degree of automation and precision of the segmentation algorithm. Next, use gray gradient changes between aortic wall and membrane to extract the aortic dissection membrane by Hessian matrix and continuity a priori model. Finally, three-dimensionally reconstruct the aorta structure by volume rendering and set the transparency and colors of the aortic dissection three-dimensional model to make the interval more intuitive and clear. Ultimately, the system can provide help for clinical diagnosis and measures.

3. Implementation Method

3.1. Aorta Segmentation. Due to the complexity of the three-dimensional structure of the aorta, especially the aortic arch and descending aorta shape big differences, we need to set up aortic arch and descending aorta training set, respectively. In this paper, the training set of the aortic arch and descending aorta are composed of a quantity of samples extracted by equal interval from the aortic dissection patients' CT sequence which contain all the changes in the shape. The two training sets can reflect all the patterns of shape change. We select multiple typical aortic dissection patients' CT images from the Tianjin chest hospital to build the aortic arch and descending aorta model library; different patients have different aortic CT sequence numbers, in the range of about 700~1000 on average. In order to model a shape, we represent it by a set of landmark points. This must be done for each shape in the training set and must be done correctly and accurately [5]. The number of feature points on each CT image must be consistent. The labeling is important, and each label represents a particular part of the aorta or its boundary [6, 7]. In general, we choose points marking parts of the object with particular application-dependent significance, points marking application-dependent things, and other points that can be interpolated from points of types above [8]. Aortic arch and descending aorta landmark points obtained by the expert are shown in Figure 3, and after marking feature points in each CT image we pick up the aortic arch and descending aorta training sample sets are derived, respectively.

According to landmark points, we build aorta big data model library. There is a big deviation in shape and position of different samples, so before modeling, we need to normalized registration shape vectors of two training sets, so as to realize the model building of the aorta [9]. Firstly, select one CT image, respectively, as a benchmark sample from two groups of big data training set, and then scale, rotate, and translate other samples in the library to align them with the benchmark sample; while all the differences between those samples and the reference sample are less than the setting threshold, registration is completed. Registration results of two groups of training set, respectively, are shown in Figure 4, and the horizontal ordinate represents the pixel coordinates in CT image.

Aortic model building is a high-dimensional data processing. In order to simplify calculation process, this paper applies Principal Component Analysis (PCA) approach to reduce dimensionality of the training set which involves all shape vectors to determine the main components [10]. The mean shape vector of the aortic training sample set is calculated by

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i, \qquad X_i = \left(x_{i1}, y_{i1}, x_{i2}, y_{i2}, \ldots, x_{in}, y_{in}\right)^T, \tag{1}$$

where N is the number of the samples of the training set; X_i represents the shape vector by stacking n landmark points; then calculate the covariance matrix S by

$$S = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)\left(X_i - \overline{X}\right)^T. \tag{2}$$

Calculate the eigenvectors ϕ_i corresponding to the eigenvalues λ_i, arrange the eigenvalues in descending order, and choose k eigenvectors which are relative to k bigger eigenvalues, several changing patterns that describe the observed

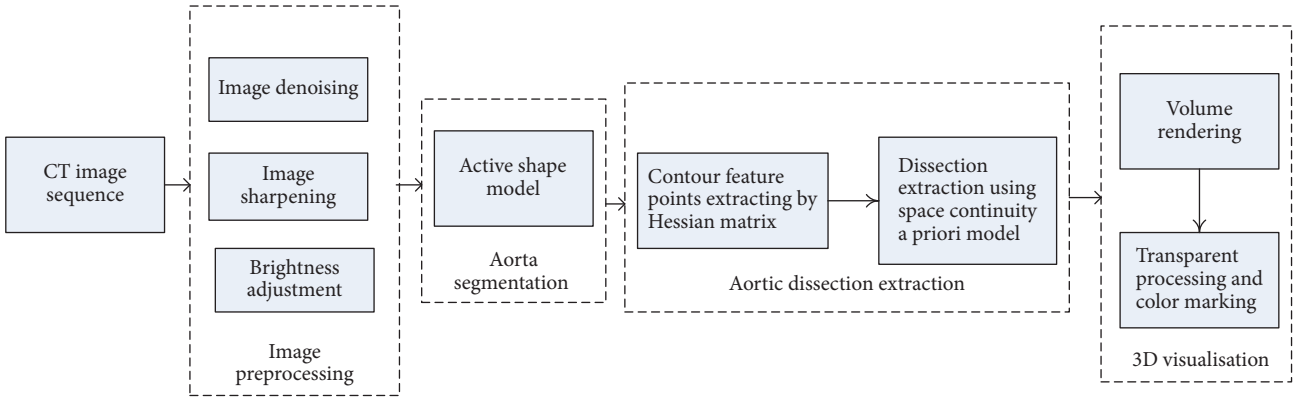
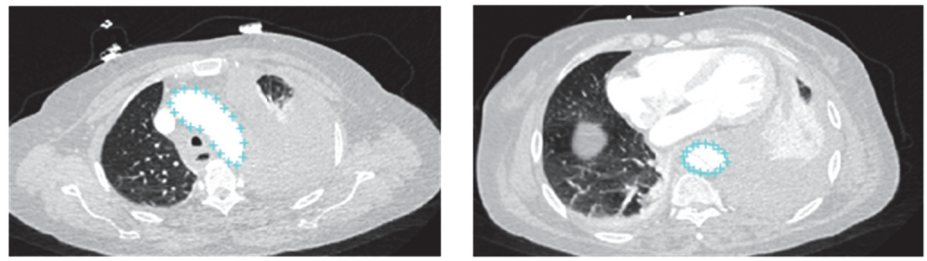


Figure 2: The framework of the three-dimensional reconstruction system.



(a) Aortic arch landmark points (b) Descending aorta landmark points

Figure 3: Landmark points marking.

variation of the training set; at the same time, k eigenvalues need to satisfy the formula as follows:

$$\sum_{i=1}^{k}\lambda_i \geq \rho\sum_{i=1}^{2n}\lambda_i, \tag{3}$$

where $\rho = 0.72$ [11]; the search accuracy can reach the highest by the principal components of 72%, and for the general value of 98%, too many constraints will appear when searching. Ultimately, the shape model can be approximated by

$$X = \overline{X} + \Phi B, \tag{4}$$

where Φ is the first k eigenvectors, $\Phi = (\phi_1, \phi_2, \ldots, \phi_k)$, and B is the projection coefficient on the principal component of the shape vector which is calculated as $B = (b_1, b_2, \ldots, b_k)^T$; suitable limits of b_i are typically determined in the range of $(-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i})$, to make sure the active shape model changes within a small range.

The gray-scale texture model is established to match and search target contour at the time that the shape statistical model is built. In this paper, the gray-scale texture model for each landmark is carried out by putting the point as the center of the square of gray sampling, avoiding the traditional method using only the vertical direction information, leading to incomplete search and error convergence. Each landmark gray vector g_{ji} mean and covariance of vector in the training sample can be approximated by

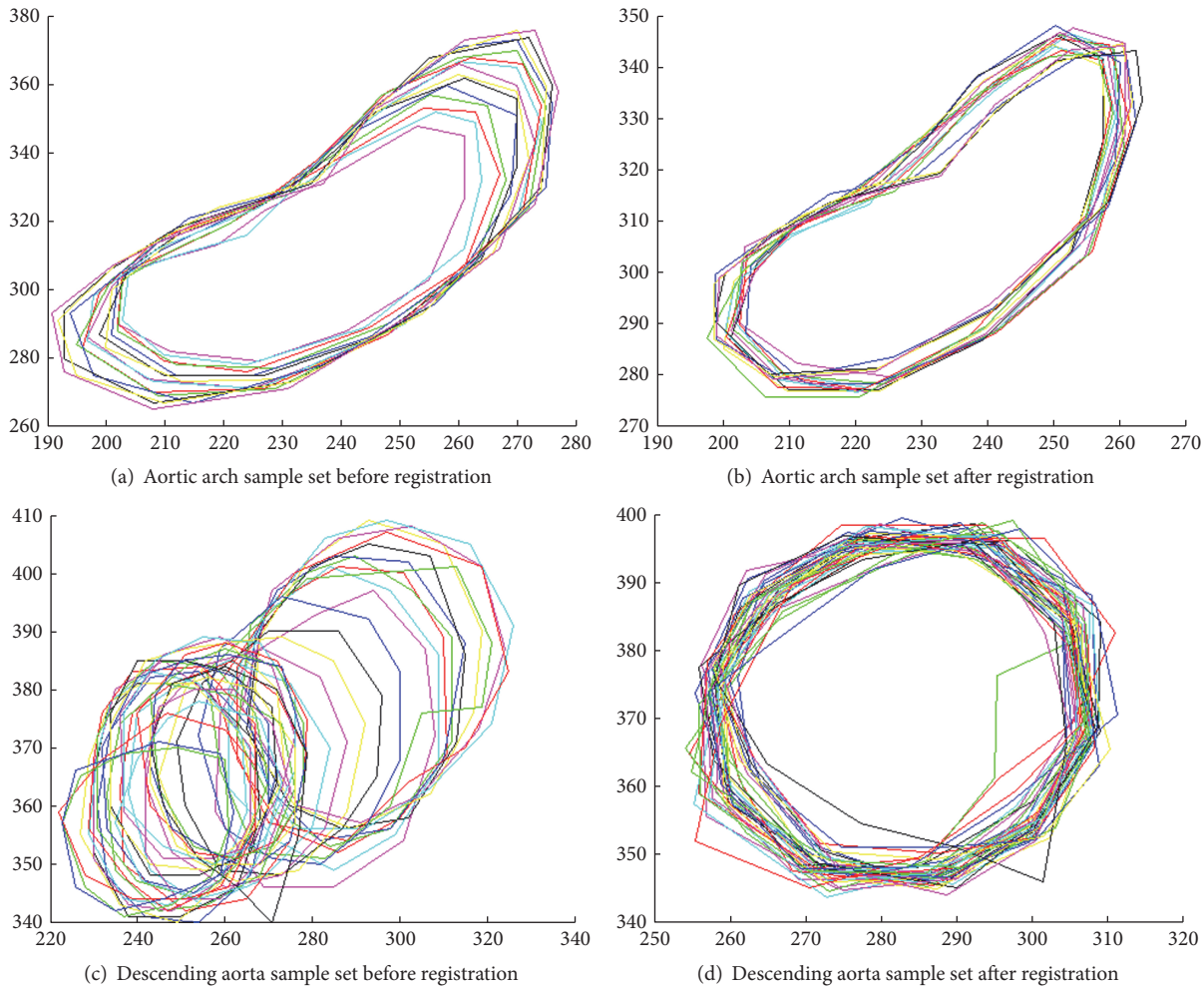
$$\begin{aligned}\overline{g}_j &= \frac{1}{N}\sum_{i=1}^{N} g_{ji}, \quad j = 1, 2, \ldots, n,\\ S_j &= \frac{1}{N}\sum_{i=1}^{N}\left(g_{ji} - \overline{g}_j\right)\left(g_{ji} - \overline{g}_j\right)^T, \quad j = 1, 2, \ldots, n.\end{aligned} \tag{5}$$

We minimize the Mahalanobis distance between a new profile and the model in the subsequent matching search process as a standard; the matching function is expressed as

$$f\left(G_j\right) = \left(G_j - \overline{g}_j\right)^T S_j^{-1}\left(G_j - \overline{g}_j\right), \tag{6}$$

where G_j is the gray vector of the jth feature point of the matching image; we get the best matching point when the matching function takes the minimum, and when we find all the matching points, the new profile is obtained.

In order to improve the efficiency and robustness of the model, multiresolution search strategy is adopted [12]. 1 mm thickness at the time of CT scanning is used; therefore a complete set of aorta CT image sequences can be seen continuous, and based on CT image sequence space continuity, by adjusting the initial shape parameters to make the contour close to the target area, the target contour parameters, which are derived from the former image, are used to adjust the current layer, by repeating the above operation between



(a) Aortic arch sample set before registration

(b) Aortic arch sample set after registration

(c) Descending aorta sample set before registration

(d) Descending aorta sample set after registration

Figure 4: Registration results.

layers. High efficient and accurate segmentation of CT images can be realized; the algorithm is of high degree automation. Figure 5 shows the final segmentation results after iteration convergence.

3.2. Aortic Dissection Extraction and Three-Dimensional Reconstruction. In this paper, we propose a detection algorithm combining Hessian matrix and spatial continuity a priori model. The Hessian matrix can be used to extract the pixels on the dissection and aortic boundary [13], the result is shown in Figure 6(a), and then use the Bayesian theory of spatial continuity model to remove other nontarget pixels, that is, only the dissection membrane pixels extracted; the result is shown in Figure 6(b). This algorithm makes full use of the continuity between the CT images of each layer and limits the offset error of the interlayer membrane to a very small range and realizes the accurate extraction of the interlayer membrane.

After the aorta segmentation and dissection membrane extraction are complete, the ray casting is carried out to reconstruct the structure. The ray casting method is a direct volume rendering algorithm based on the image sequence [14]. The specific reconstruction process is as follows.

(1) Read the three-dimensional discrete data field; set different opacity values and color values according to the size of the voxels' pixel values of vertices.

(2) A ray is emitted from each point of the screen based on the direction of the line of sight so that the rays pass through the data field space and k samples are selected equidistantly across all the rays.

(3) The opacity value and the color value of each sampling point are calculated by the trilinear interpolation algorithm using the data values of the eight vertices nearest to the sampling point.

(4) Calculate the opacity and color values for all pixels on the screen based on the cumulative order from front to back.

(5) The opacity value and the color value of each pixel obtained are projected onto the imaging screen to generate the final three-dimensional image.

(a) Descending aorta from segmentation (b) Aortic arch from segmentation

Figure 5: Final result after iteration convergence.

(a) Contour pixel feature points extraction result (b) Aortic dissection results

Figure 6: Aortic dissection extraction result.

The final 3D reconstruction of a group of aortic region is as shown in Figure 7. By setting the aorta transparency and color, the interlayer is shown in red to give the final reconstruction results, as shown in Figure 8.

4. Experimental Results and Discussions

4.1. Experimental Results. Figure 9 shows a three-dimensional model of the aortic dissection of the other two patients reconstructed by the above method. It is clear from the three-dimensional reconstructed images that the dissection membrane occurs in the entire aortic cavity in the first set of models, and the dissection membrane of the second group appears in the aortic arch away from the heart. The results of the two groups intuitively and clearly show the location and extent of the break, so the method provides much more space relationship information of the aortic dissection to the attending physician for the diagnosis of the diseases, surgery, and postoperative evaluation to provide assistance.

4.2. Segmentation Accuracy. In order to verify the algorithm for segmentation of CT image sequence aortic extraction effect, compare the aortic arch and descending aorta extraction results on this proposed method with those based on GVF snake algorithm. To test and verify the reliability of above two algorithms, we compare the extraction results by algorithms above with the manual results by doctor of rich clinical experience and calculate the overlapping rate to test the reliability of the algorithm in this paper; the extraction results by three methods are shown in Figure 10.

We can obtain that, for descending aorta, the segmentation results by both GVF snake model and the proposed segmentation algorithm are close to the real target contour. While for the aortic arch the GVF snake algorithm has certain errors, because of the absence of shape constraints, the

Table 1: Two segmentation algorithms' overlap ratio.

	GVF snake method (aortic arch)	Our proposed method (aortic arch)	GVF snake method (descending aorta)	Our proposed method (descending aorta)
Sample 1	84.52%	92.32%	89.75%	95.59%
Sample 2	86.51%	92.97%	88.16%	94.98%
Sample 3	87.63%	93.09%	90.03%	93.26%
Sample 4	89.95%	94.65%	89.80%	93.76%
Sample 5	88.02%	93.23%	88.70%	94.07%
Sample 6	88.63%	93.45%	89.04%	94.12%

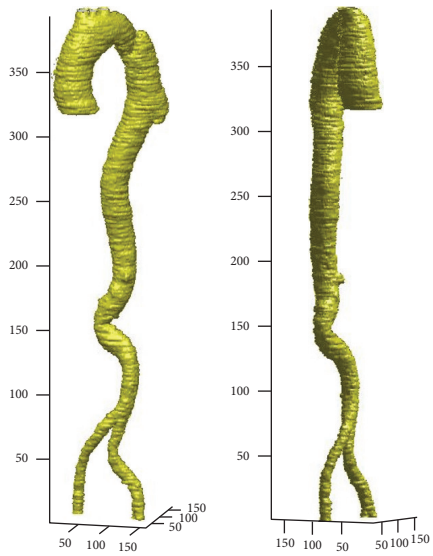


Figure 7: 3D visualization of the aorta region.

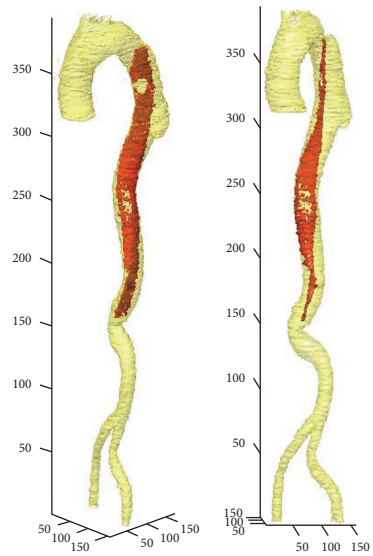


Figure 8: 3D visualization of aortic dissection.

curve is vulnerable to the interference of other organizations around the outline in the process of evolution, which is shown in Figure 11(a); curve 1 is the initial curve, and curve 2 is the end evolution of the curve. In order to avoid impact from other groups' contour in the curve evolution, the GVF snake model firstly determines the aorta interested area which involves complete aorta area but as far as possible to introduce the interference of other organizations [15, 16]. Generally speaking, fixed location areas or template matching methods are chosen, which are with low repeatability and adaptability, and the selecting accuracy directly affects the subsequent partitioning extraction results. However we can achieve the aorta automatic segmentation result after directly inputting one CT image with our proposed method. Figure 11(b) shows the iteration results by our proposed method. And this algorithm is of higher degree of automation and stronger robustness than GVF snake model.

In order to quantitatively describe the accuracy of the algorithm in this paper, using the overlap rate as evaluation indexes,

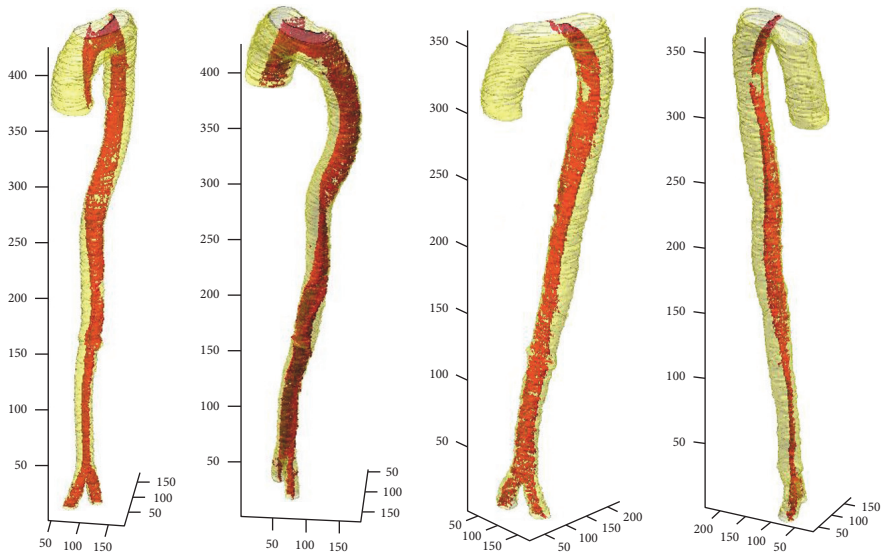
$$\text{Overlap} = \frac{A_{ab}}{A_a + A_b - A_{ab}}; \tag{7}$$

A_a and A_b, respectively, represent the proportion of the target area from the two images' segmentation; A_{ab} is the proportion of the target zone overlapping. With thoracic hospital doctors manual segmentation result as the gold standard, our experiment selects large number of images from groups of patients' CT images and then calculates the two algorithms' overlap, respectively, and some samples results of those all are displayed in Table 1.

For extraction accuracy, the proposed method is obviously of higher overlap ratio than the GVF snake method from the table (93.29% versus 87.54% on aortic arch, 94.30% versus 89.25% on descending aorta).

5. Conclusions

This paper introduces a kind of medical CT image processing method to rapidly and accurately obtain aortic dissection characteristic from huge CT image data and to three-dimensionally reconstruct the structure for doctors in clinical diagnosis. The algorithm of this paper has been improved on the basis of the traditional ASM algorithms; not only is the statistical model more accurately constructed and the accuracy of the model matching improved, but also the initialization process combined with the continuity of the



(a) The first patient's aortic dissection display (b) The second patient's aortic dissection display

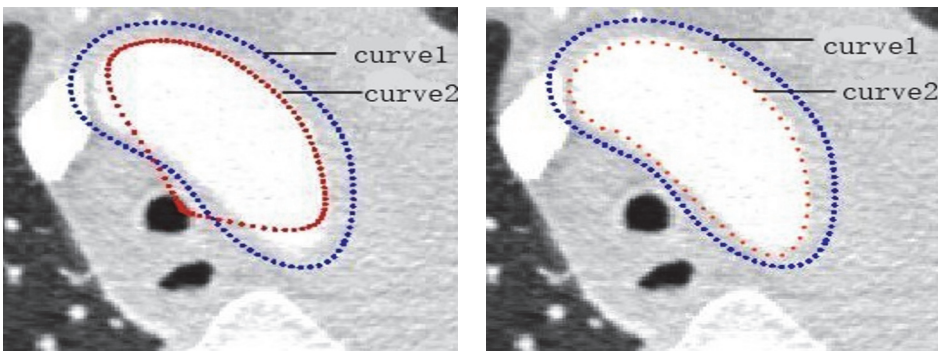
Figure 9: 3D visualization model of another two patients with aortic dissection.

(a) Aortic arch and descending aorta manual segmentation results

(b) Segmentation results by GVF snake method

(c) Segmentation results by this proposed method

Figure 10: Aortic arch and descending aorta segmentation results by three methods.



(a) Iteration result based on GVF snake method (b) Iteration result based on the proposed method

Figure 11: Iteration results based on two methods.

CT sequence is simplified and the algorithm effectiveness is improved. Compared to GVF snake algorithm, the experiment can be a strong proof that our algorithm can effectively improve the accuracy of aorta segmentation. This method of this paper can effectively make up for the inadequacy of existing hospital equipment function, so that the attending physicians and patients can deeply understand and grasp the state of the illness. At the same time, for other cardiovascular diseases' diagnosis and treatment, this method is of great significance.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (NSFC) (61405143, 61373104, and 61402330) and Key Projects in the National Science and Technology Pillar Program (2013BAF06B00) under Tianjin Key Laboratory of Optoelectronic Detection Technology and Systems, Tianjin Polytechnic University, China.

References

[1] N. Fetnaci, P. Lubniewski, B. Miguel, and C. Lohou, "3D segmentation of the true and false lumens on CT aortic dissection images," in *Proceedings of the Three-Dimensional Image Processing (3DIP) and Applications 2013*, Burlingame, CA, USA, February 2013.

[2] C. Lohou, N. Fetnaci, P. Lubniewski, B. Miguel, P. Chabrot, and L. Sarry, "Intimai flap segmentation on CTA aortic dissection images based on mathematical morphology," in *Proceedings of the 2013 6th International Conference on Biomedical Engineering and Informatics, BMEI 2013*, pp. 143–147, Hangzhou, China, December 2013.

[3] M. Deeley A, A. Chen, R. Datteri et al., "Comparison of manual and automatic segmentation methods for brain structures in the presence of space-occupying lesions: a multi-expert study," *Physics in Medicine and Biology*, vol. 56, no. 14, p. 4557, 2011.

[4] T. F. Cootes and C. J. Taylor, "Statistical models of appearance for medical image analysis and computer vision," in *Proceedings of the Medical Imaging 2001 Image Processing*, pp. 236–248, San Diego, Calif, USA, February 2001.

[5] C. Je, W. Jo, and H.-M. Park, "Mouth map-assisted localisation of ASM lip landmarks," *Imaging Science Journal*, vol. 64, no. 8, pp. 419–424, 2016.

[6] Z. Zheng, J. Jiong, D. Chunjiang, X. Liu, and J. Yang, "Facial feature localization based on an improved active shape model," *Information Sciences*, vol. 178, no. 9, pp. 2215–2223, 2008.

[7] G. Gill, M. Toews, and R. R. Beichel, "Robust initialization of active shape models for lung segmentation in CT scans: A feature-based atlas approach," *International Journal of Biomedical Imaging*, vol. 2014, Article ID 479154, 2014.

[8] T. Cootes, C. Taylor J, and A. Lanitis, "Active shape models: evaluation of a multi-resolution method for improving image search," in *Proceedings of the British Machine Vision Conference*, pp. 327–336, 2010.

[9] H. Sun, X. Lu, and H. Liu, "A segmentation method of aorta in ct image," *Modern Scientific Instruments*, no. 02, pp. 45–48, 2013.

[10] J. Wang and C. Shi, "Automatic construction of statistical shape models using deformable simplex meshes with vector field convolution energy," *Biomedical Engineering Online*, vol. 16, article 49, no. 1, 2017.

[11] S. Kurugol, R. San Jose Estepar, J. Ross, and G. R. Washko, "Aorta segmentation with a 3D level set approach and quantification of aortic calcifications in non-contrast chest CT," in *Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '12)*, pp. 2343–2346, IEEE, San Diego, Calif, USA, September 2012.

[12] C.-Y. Tsai, C.-H. Huang, and A.-H. Tsao, "Graphics processing unit-accelerated multi-resolution exhaustive search algorithm for real-time keypoint descriptor matching in high-dimensional spaces," *IET Computer Vision*, vol. 10, no. 3, pp. 212–219, 2016.

[13] R. Lakemond, C. Fookes, and S. Sridharan, "Affine adaptation of local image features using the Hessian matrix," in *Proceedings of the 6th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2009*, pp. 496–501, Genova, Italy, September 2009.

[14] J. Sun, H. Li, P. Gao, and L. Wu, "Research on high efficient ray casting algorithm based on VTK," in *Proceedings of the 7th International Conference on Information Technology in Medicine and Education, ITME 2015*, pp. 212–214, Huangshan, China, November 2015.

[15] T. Guan, D. Zhou, and Y. Liu, "Accurate segmentation of partially overlapping cervical cells based on dynamic sparse contour searching and GVF snake mode," *Journal of Biomedical and Health Informatics*, vol. 19, no. 4, pp. 1494–1504, 2015.

[16] F. Zhang, X. Zhang, K. Cao et al., "Contour extraction of gait recognition based on improved GVF Snake model," *Computers & Electrical Engineering*, vol. 38, no. 4, pp. 882–890, 2012.

Hindawi
International Journal of Digital Multimedia Broadcasting
Volume 2017, Article ID 4815932, 7 pages
https://doi.org/10.1155/2017/4815932

Research Article

Research of Simulation in Character Animation Based on Physics Engine

Yang Yu,[1] Jucheng Yang,[1] Xiaofei Zan,[2] Jiangang Huang,[1] and Xiangbo Zhang[1]

[1]*College of Computer Science and Information Engineering, Tianjin University of Science and Technology, Tianjin, China*
[2]*School of Computer and Information Technology, Beijing Jiaotong University, Beijing, China*

Correspondence should be addressed to Jiangang Huang; huangjg@tust.edu.cn

Received 28 February 2017; Accepted 17 May 2017; Published 12 July 2017

Academic Editor: Jenq-Neng Hwang

Computer 3D character animation essentially is a product, which is combined with computer graphics and robotics, physics, mathematics, and the arts. It is based on computer hardware and graphics algorithms and related sciences rapidly developed new technologies. At present, the mainstream character animation technology is based on the artificial production of key technologies and capture frames based on the motion capture device technology. 3D character animation is widely used not only in the production of film, animation, and other commercial areas but also in virtual reality, computer-aided education, flight simulation, engineering simulation, military simulation, and other fields. In this paper, we try to study physics based character animation to solve these problems such as poor real-time interaction that appears in the character, low utilization rate, and complex production. The paper deeply studied the kinematics, dynamics technology, and production technology based on the motion data. At the same time, it analyzed ODE, PhysX, Bullet, and other variety of mainstream physics engines and studied OBB hierarchy bounding box tree, AABB hierarchical tree, and other collision detection algorithms. Finally, character animation based on ODE is implemented, which is simulation of the motion and collision process of a tricycle.

1. Research Background and Significance

The mainstream animation production technologies are key frame technology based on artificial production and motion capture technology based on capture equipment. Key frame technology based on artificial production is the original animation combination technology. It depends on the key frame production and the interpolation between key frames to drive the continuous animation playing. The technology allows designers to create animation by their own will with the greatest degree of freedom, but the products of the technology are not very good in terms of natural performance; moreover, the process of production is quite complicated [1]. Since the motion capture equipment appeared, it can be used to capture the motion data as the information source of animation production, so that the efficiency of producing character animation and the reality sense of character animation can be greatly improved. However, since the technology needs the genuine characters to take part in the process of capturing the motion information, its application scope becomes quite limited. When using this technology, it is not allowed for the genuine actors to take some dangerous actions. Furthermore, the high cost of this kind of equipment will also restrict the wide use of the technology.

In the background of growing industrial demand, increased user experience requirement, and fast developing of computer-related technologies, the above-mentioned two kinds of common traditional animation production methods are becoming less and less possible to satisfy the requirements of design and use. The root cause is that the two kinds of technologies use the offline way during the process of animation combination and storage and then for the post-production use; thus, the animation will be charged with the corresponding huge data storage space considering its high quality and long duration. Moreover, the biggest issue of this method is the poor rate in terms of real-time and self-adaption performance, so that the animation, which has cost plenty of time and money, can be just used for the specific

circumstance for once. All of these issues do not fit into the current trend of fast increase of computer processing and user experience [2].

This article, in order to deal with the issues of poor real-time performance, low utilization rate, and complicated procedure, existing during the process of producing character animation, does research on the character animation simulation method based on physics engine and uses an example to implement the method.

2. Foundation of Related Researches

Nowadays, character animation is always the focus and difficulty in the computer graphics research. In the research, there are various animation production methods, which can be categorized as follows [3]:

(i) The animation production method based on kinematics

(ii) The animation production method based on dynamics

(iii) The animation production method based on motion data

A brief introduction of the above three kinds of character animation is next.

2.1. Principle of Character Animation Production Based on Kinematics. Kinematics is a very important branch in physics mechanics. During the research, the object is abstracted as a rigid body model and a particle model [4]. These models are used to study the motion of objects. The key point of the rigid body model is to research the rigid body's angular velocity, rotary motion, linear velocity, and so forth. The focus of the research on the particle model is on the velocity, acceleration, deceleration, and motion equation of the particle in the specified reference system. In addition, it is not necessary to consider the influence factors of dynamics, such as quality and force.

The kinematics method is divided into forward kinematics and inverse kinematics.

2.1.1. Forward Kinematics Method [5]. In a simple kinematic system, there are often two ends. They are free end and fixed end. Forward kinematics describes the motion of the free end with the fixed end of the system as the starting point. Forward kinematics method: the first step is to establish the hierarchical structure of the object and then to establish the rotation and position of each object in the hierarchy.

2.1.2. Inverse Kinematics Method. Another simulation technique based on kinematics is the inverse kinematics method [5]. Inverse kinematics is just opposite to forward kinematics. The inverse kinematics is the motion of which the fixed end drove by the free end.

Use the following formula to express the basic principles of the inverse kinematics [6]:

$$\Delta X = J\Delta\theta. \tag{1}$$

Formula (1) is the basic formula of the inverse kinematics, in which ΔX is the displacement of the end of the joint chain, $\Delta\theta$ is the rotate angle of the end of the joint chain, and J is the Jacobian matrix. By the formula above, the displacement of joint is equal to a product, the rotation angle of joint chain ends, and the Jacobian matrix. But formula (2) is frequently used in inverse kinematics [6]:

$$\Delta\theta = J^{+}\Delta X + \alpha\left(I - J^{+}J\right)\Delta z. \tag{2}$$

In the formula, ΔX and $\Delta\theta$ are similarly found in formula (1). J^{+} is pseudo-inverse matrix of the Jacobian matrix in linear algebra. I is a unit matrix. α is optimization constant. Δz is the energy consumption of joints in minimizing motion. The inverse kinematics is calculated from the end of the entire joint chain. Compared with the forward kinematics, inverse kinematics is more suitable for the creation of more complex motion and easier to enable the flip of the object in the animation. However, the inverse kinematics is not universal. Some problems which are suitable for forward kinematics are not suitable for inverse kinematics.

2.2. Principle of Animation Production Based on Dynamics. Dynamics is also a very important branch of physics, which researches the relationship between force and motion. The basic theory of kinetic research is Newton's law of motion. The object of the dynamics study is a macroscopic object that is far lower than the speed of light. The animation produced by this method has a strong physical reality, because it is based on the laws of physics.

In the simulation of character animation, kinematics method is to calculate the position and rotate angle of each joint for the characters, and the key frames are composed of two kinds of data.

In the simulation of character animation, the dynamic method first gives the object to the physical properties and exerts various forces and then uses Newton's law of motion to calculate the state of the character. The essence of this method is to calculate the acceleration of an object by force. In dynamics, the type of force can be divided into two types: point force and field force [7].

In addition to the force, the collision is also an important factor affecting the movement of the character. In the real world, when two objects collide, they collide, but they do not penetrate each other. When the collision occurs, the interaction force will change the original state of the two objects, such as position, trajectory, and direction. Therefore, in the simulation of the character animation, if the collision is between objects, the original motion state of the character will be changed because of the interaction force.

2.3. Principle of Character Animation Production Based on Motion Data

2.3.1. Motion Capture. The animation methods mentioned above are realized by the use of mechanics, biomechanics, and daily life experience. In motion capture technology, data is originated from the real actor performance, through motion capture equipment to obtain action data. Because the processed data is produced by human beings, the motion capture technology can make the animation more real and natural. There was no artificial data from capture and processing. But, motion capture equipment is very expensive and complicated to operate. During the data collection, the motion model's limb, body, and head commonly set sensors point, and those collective data are a very important part in whole motion capture technology [8].

The dynamic and kinematics techniques introduced in this paper are used to animate the physical laws of motion. However, the complexity of human motion with physical laws, especially in terms of human motion coordination mechanism, makes simulation work very difficult, thus leading to the produced animation being not realistic and the lack of rich detailed information of actual human body movement. In recent years, the motion capture technology has entered the line of sight; it overcomes the shortcomings of the above two methods of animation and has become one of the most promising technologies in human animation.

2.3.2. Motion Blend. The motion blending technology is a technology based on special motion database and algorithm. To use this technique to deal with animation, we generally need to use interpolation algorithm and interpolation parameters.

The main advantage of motion blend technology is its low computational complexity. And it is based on motion capture data, to maintain the original dynamic motion. Its main drawback is that, for example, tension, worry, and other acts are too dependent on the existing operating data, and the amount of data increases. If a character is walking in a gait that is not too tired or tired, it is necessary to have a normal walking gait data and also to have a very tired walking data. In the interpolation process, the producers usually deal with the transition and transformation between the collection of parameters by the persistent and linear default views [9].

2.3.3. Motion Deformation. Motion deformation is the technology to modify and use the existing motion information. Popoivc and Witkin use interactive methods to shift and scale the selected key frames to modify the existing data. Compared with the trajectory in the time domain, Williams and Bruderlin are used to adjust, modify, and reuse the existing animation data in the frequency domain. Compared with moving animation technology and process animation technology, the advantage of motion mixing technology is that it can use the existing real capture data for animation generation. Moreover, this technology gives the animation creator and key frame animation compatible with a tool that will be very convenient to deal with.

Compared with kinematics animation technology and process animation technology, the advantage of the motion blending technology, which belongs to the process animation technology, is that it can make use of the existing real capture data to generate animation [10]. Moreover, this technology is compatible with key frame animation technology.

3. The Physics Engine Related Research

Generally, there are two types of physics engines: open source physics engine and commercial physics engine.

Commonly used professional physics engines include ODE, PhysX, Bullet, Havork, Newton, and Vortex. These physics engines are widely used in the domains as graphics modeling tools, game development, virtual reality systems, and so on.

Collision detection is the technical core of physics engine, so most of the research on physics engine is the research on the collision detection algorithm. In real life, some reactions occur naturally at and after the collision of objects [11]. However, in the virtual world, the collision issue needs some algorithms to be solved. Collision detection is the important part of the domains like Robot, VR, and so on. Its main purpose is to determine whether two or more objects will penetrate or contact. The research of collision detection is, through a certain algorithm, to enable two or more objects not to occupy the same spatial area at the same time, without damaging the objects.

At present, there are two major categories of research on collision detection worldwide: spatial decomposition method and bounding box method. The bounding box method simulates a complexly shaped object into an approximate simple three-dimensional space. The hierarchical bounding box method of the bounding box method is suitable for the collision detection in complex environment.

The ODE (Open Dynamics Engine) is an open source library of rigid body dynamics that is often used in the simulation of joint connections. ODE can be applied to simulate the moving objects in the VR environment, legged creatures, and vehicles on the ground. It is robust, flexible, and fast. Flexibility is reflected in the portability; Unix/Linux, Windows, and MacOS can be used. And it has a built-in collision detection system.

4. Research and Design of Character Animation Simulation Based on ODE

4.1. Dynamics Modeling of Simulated Vehicle Based on ODE. In ODE, there is a special type of joint as shown in Figure 1.

Hinge-2 joint is made up of two noncoaxial hinges. The utility model comprises two rotating shafts and a supporting point. Among them, Axis 2 can only rotate around the axis; however, Axis 1 not only can rotate around the axis but also can limit the scope of its rotation.

The car model consists of body and wheels. According to the joint type of Hinge-2, Body 1 can represent the body and Body 2 can represent the wheel; similarly, Axis 1 can represent the steering wheel of the wheel and Axis 2 can represent the

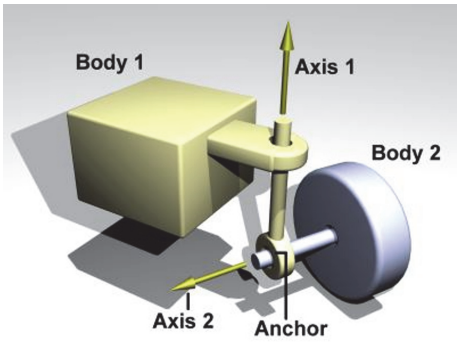


Figure 1: Hinge-2 joints [20].

wheel rolling shaft. In the dynamics design of the ODE system, ODE rigid object with the quality and position is generated, and ODE geometry corresponding to the rigid body is created at the same time. Finally, the body and the wheels are connected by Hinge-2 joint. A simple car model has been built.

4.2. Simulation of Character Collision Based on ODE. The simulation process of character collision based on ODE is as follows.

4.2.1. Spatial Initialization in Collision Simulation. At this stage, it is necessary to set the type of collision space and the relevant properties of the contact surface, such as flexibility, elasticity, and friction.

4.2.2. Data Initialization in Collision Simulation. At this stage, it is necessary to do some work on the initialization and cleanup of the data before the collision simulation [12].

Collision Data Cleaning. Clear all kinds of collision information, including clearing collision data, clearing contact group, and destroying the current collision space. Then, reset all the properties of the contact surface to comply with the needs of the user.

According to user needs, create four-fork tree space, hash space, and simple space. These three new types of collision space use different collision detection algorithms and different data structures. Different geometries are stored using different data structures.

Collision Data Detection. Collision data is a custom variable that stores all the data needed for a collision geometry. The character object is the source of the data. Prior to the crash simulation, you need to set up a collision data for each collision object and add the data to a list of collisions. Finally, the collision information is transferred to the collision function when the collision simulation is performed.

4.2.3. Collision Simulation Start. The simulation process takes the time step as a unit. With the ODE simulator, simulation to complete a time step for the unit collision effect needs to use the current state before the impact and role of the data and then the collision effect exerted on the role of the body [13, 14]. This link needs to call the existing space collision function in ODE. Finally, according to the data of the contact nodes generated by the two contact rigid bodies in the collision, the ODE completes the collision [15]. The space collision function is a very important function in this link. Its function prototype is

```
Void dSpaceCollide (dSpaceID space,
void *data, dNearCallback *callback).
```

This function determines which geometries are likely to occur in space and then calls the callback function to handle the pair of geometries and finally completes the collision.

Callback's function type is dNearCallback, defined as follows:

```
Typedef void dNearCallback (void
*data, dGeomID o1, dGeomID o2)
```

Among them, data is user-defined parameters, directly from the dSpaceCollide pass. A pair of geometries that detect a possible collision are represented by parameters $o1$ and $o2$ [16]. This callback function then uses dCollide to generate the contact point between the geometries. Then, these contact points will be added to the simulation process as a contact joint.

The following operation will occur inside the callback [17]:

(i) Use the dCollide function to generate the selected contact points between the geometries, while returning the number of contact points.

(ii) Based on all the collision data in the previous link, for each contact point, two eligible collision data points are searched.

(iii) ODE completes the collision effect according to the contact joint information contained in these two collisions and shows the picture again.

4.2.4. Collision Simulation End. With the end of the ODE simulation, the collision simulation also acts as a driving force to stop the character of force. This link will do some related data cleanup work, which is conducive to the next simulation.

5. Implementation and Testing of Character Animation Simulation

5.1. Creating an Example in the Character Animation. Open the physics engine software and edit and run the C language program to get the scenario shown in Figure 2.

In Figure 2, you can see the basic elements of the scene, including the sky, the ground, the car, and the obstacles.

The simulation example consists of simple elements: the power of vehicle comes from the motor; the power output and the steering are set in the front wheel; the wheel and the vehicle body joint through the hinge; the vehicle may have interaction forces when touching any object of the scenario; there is a collision detection device in the vehicle.

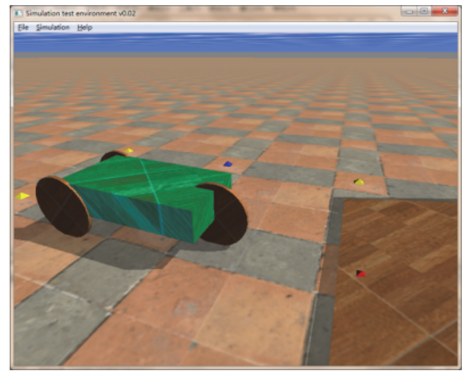


Figure 2: Scenario picture.

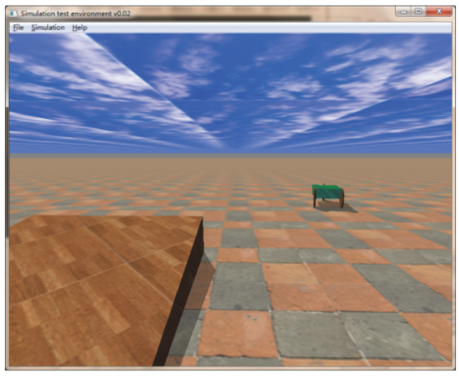


Figure 3: Forward, backward, acceleration, and deceleration of the vehicle model.

5.2. Testing. By setting the power of the vehicle, the interaction force between the vehicle and the objects in the scenario, as well as the ground slip rate, the simulation of the vehicle motion, and collision process can be implemented.

The performance testing is shown in Figures 3, 4, and 5.

In the course of the testing, change the vehicle or the various parameters of the surrounding environment and observe whether the character change in the pictures conforms to the kinetic principle.

5.2.1. Friction Testing. When the vehicle has contact with the ground, change the Coulomb friction coefficient mu from the original dInfinity (there will not be sliding contact) to 0 (no friction contact). This coefficient must be set. And the testing result is that the wheel is idling. At this time, the front wheel as the power output can only be idling, since there is no friction, so the vehicle cannot move [18].

5.2.2. Penetration Testing. When the vehicle is in contact with the ground, soft_cfm is the contact normal flexibility parameter, which is used to adjust the flexibility of the object. Enlarging the soft_cfm value, from the original 0.3 to 1, indicates that there will be very obvious object penetration phenomenon; the testing result is shown in Figure 6; there appears very obvious penetration phenomenon between the wheel and the ground [18].

5.2.3. Slide Coefficient Testing. Set the FDS (Force Dependent Slip factor) to zero. After the vehicle model rolls over, continue to give power to the vehicle, and the vehicle has no side slide but has been static. (FDS indicates the lateral slide phenomenon due to the lateral external force when the object is moving, and the sliding distance is proportional to the moving speed of the object and the external force, so FDS is also called lateral force slide coefficient [18, 19].)

If the FDS slide coefficient of friction is set to 0.1, continue to give power to the vehicle after its rollover, and the vehicle can still move laterally. The testing result is shown in Figure 7.

When the world is created, the z-axis direction of gravity acceleration that is set to 0 can be obtained as in Figure 8.

From the test results, due to the acceleration of gravity of 0, there is simulation of the suspension in the vehicle.

Through the simulation of the car and the surrounding environment of various parameters set, you can carry out a variety of physics engine-based animation simulation.

6. Summary

Regarding the research on the simulation of character animation, there have been many achievements at home and abroad. On the basis of studying and researching the existing scientific research achievements, this article has done the following works: in-depth research on the mainstream computer character animation production technologies; in-depth analysis and research on the mainstream physics engines; in-depth research on the related collision detection algorithms; designing and building the simulation architecture, to make the character animation simulation feasible by using ODE engine; building the simulation experiment platform and successfully operating examples

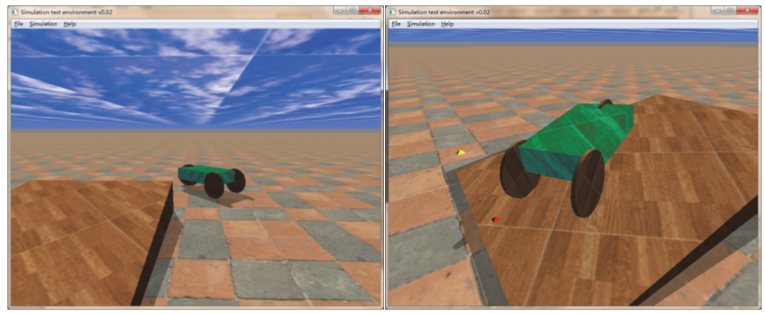


Figure 4: Turning and climbing of the vehicle model.

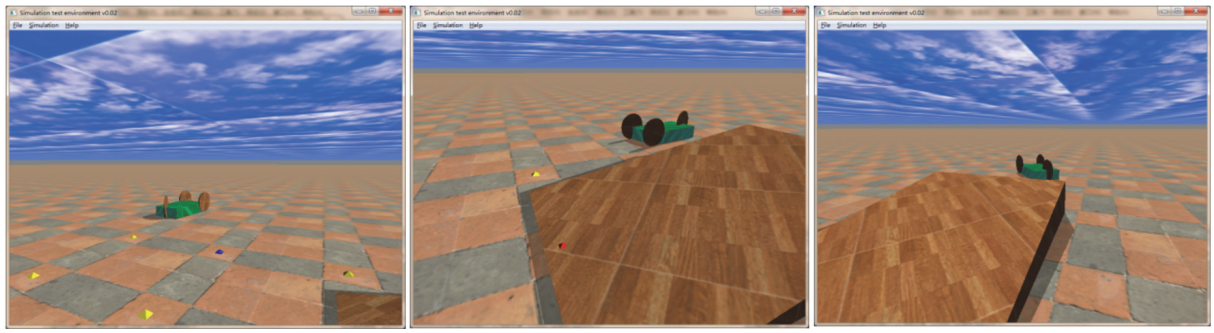


Figure 5: The vehicle model has high-speed collision with obstacle and rollover, due to high-speed turning and rushing out of the ramp.

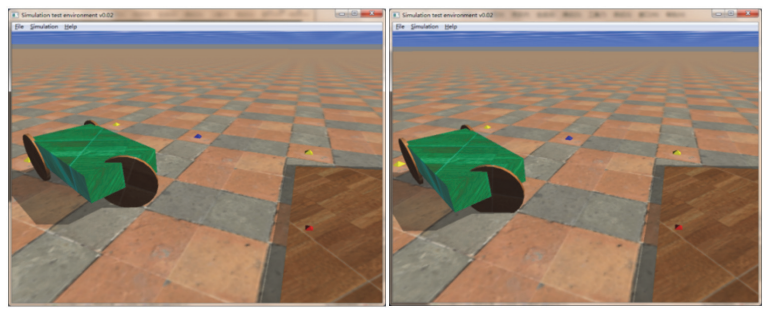


Figure 6: Comparison between preadjusted and adjusted penetration effects.

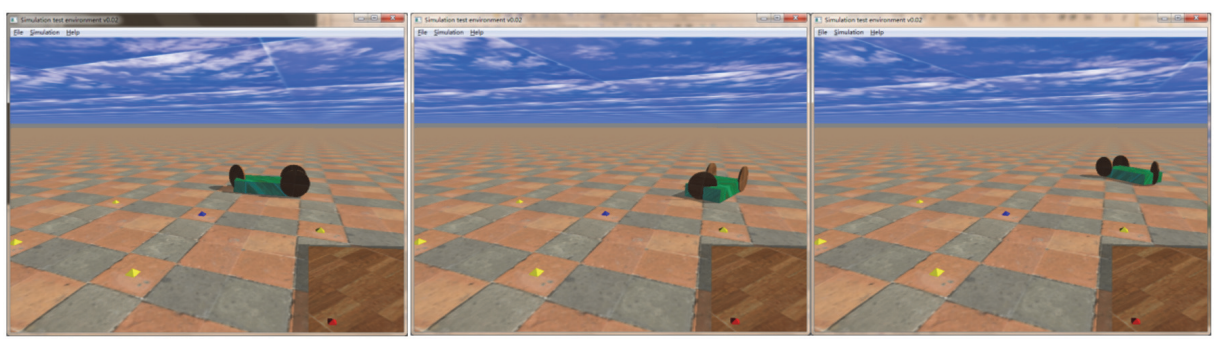


Figure 7: Horizontal movement.

to implement the ODE-based character animation simulation; and carrying out the test and analysis with different parameters.

The focus for next research includes the following: from the vehicle motion model upgrading to the more complex human body motion model, simulating the more complex and precise movements of human body; establishing different controllers for different character models and motion states and implementing more efficient and more realistic simulation through more intelligent controllers; enabling the natural switch-off states of character through animation blending technology.

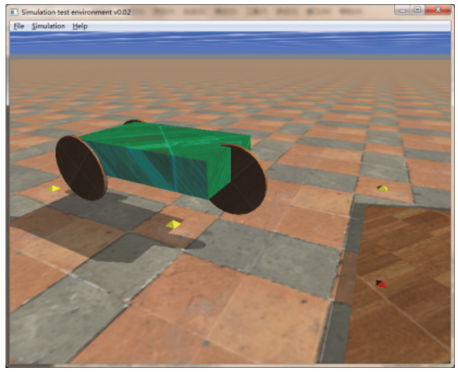


Figure 8: Vehicle suspension.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] B. Dai, *Research and Implementation of Role Based Animation Simulation*, University of Electronic Science and Technology of China, 2012.

[2] C. K. Liu, A. Hertzmann, and Z. Popović, "Composition of complex optimal multi-character motions," in *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '06)*, pp. 215–222, ACM, Vienna, Austria, September 2006.

[3] H. Duan, *Research and Application of Key Technologies of Physics Engine in Virtual Reality*, Shandong University of Science and Technology, 2010.

[4] http://baike.baidu.com/view/85895.htm.

[5] S. Gao, *Design and Implementation of Driving Simulation System Based on OGRE and ODE*, Wuhan University of Technology, 2006.

[6] V. B. Zordan and J. K. Hodgins, "Motion capture-driven simulations that hit and react," in *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '02)*, pp. 89–96, ACM, July 2002.

[7] Z. Wang, H. Jiazhen, and H. Yang, "Summary of research on the collision detection problem," *Journal of Software*, no. 5, 1999.

[8] T. Larsson and T. A. Moiler, "Collision detection for continuously deforming bodies," in *Proceedings of the Eurographi cs'*, 2001.

[9] Z. Popović, "Controlling physics in realistic character animation," *Communications of the ACM*, vol. 43, no. 7, pp. 51–58, 2000.

[10] M. Zhang, *Research on Dynamic Fault Diagnosis Method for Vehicle Mounted Track*, North Central University, 2014.

[11] T. Zhang, *Research on Automobile Virtual Driving Training System Based on Quest 3D*, Henan Polytechnic University, 2012.

[12] X. Yang, *Improvement and Optimization of Collision Detection Module for Space Desktop System*, Capital Normal University, 2011.

[13] T. Wang, "Adaptive stochastic collision detection between deformable objects using particle swarm optimization," in *Proceedings of the EvoWorkshops (EvoIASP '06)*, 2006.

[14] F. A. Madera, A. M. Day, and S. D. Laycock, "A hybrid bounding volume algorithm to detect collisions between deformable objects," in *Proceedings of the 2nd International Conferences on Advances in Computer-Human Interactions (ACHI '09)*, pp. 136–141, IEEE, Cancún, Mexico, February 2009.

[15] Y. Zou, G. Ding, M. Xu, and Y. He, "Overview of real time collision detection algorithm," *Computer Application Research*, no. 1, 2008.

[16] J. T. Klosowski, M. Held, J. S. B. Mitchell, H. Sowizral, and K. Zikan, "Efficient collision detection using bounding volume hierarchies of k-DOPs," *IEEE Transactions on Visualization and Computer Graphics*, vol. 4, no. 1, pp. 21–36, 1998.

[17] R. Smith, "Open Dynamics Engine v0.5 User Guide," 2006, http://www.ode.org/ode-latest-userguide.html.

[18] T. Geijtenbeek and N. Pronost, "Interactive Character Animation using Simulated Physics," Games and Virtual Worlds, 2011.

[19] M. Wang, *Research and Practice of Virtual Reality Art in Engineering*, Harbin Engineering University, 2010.

[20] P. Jiménez, F. Thomas, and C. Torras, "3D collision detection: A survey," *Computers & Graphics*, vol. 25, no. 2, pp. 269–285, 2001.

Hindawi
International Journal of Digital Multimedia Broadcasting
Volume 2017, Article ID 8594863, 5 pages
https://doi.org/10.1155/2017/8594863

Research Article

Prior Knowledge-Based Event Network for Chinese Text

Yunyu Shi,[1] Jianfang Shan,[2] Xiang Liu,[1] and Yongxiang Xia[1]

[1]*School of Electronic & Electric Engineering, Shanghai University of Engineering Science, 333 Longteng Road, Songjiang District, Shanghai, China*
[2]*School of Information, Qilu University of Technology, 3501 Daxue Road, Changqing District, Jinan, China*

Correspondence should be addressed to Jianfang Shan; sjfshan@163.com

Received 13 January 2017; Accepted 7 May 2017; Published 4 June 2017

Academic Editor: Hyo-Jong Lee

Text representation is a basic issue of text information processing and event plays an important role in text understanding; both attract the attention of scholars. The event network conceals lexical relations in events, and its edges express logical relations between events in document. However, the events and relations are extracted from event-annotated text, which makes it hard for large-scale text automatic processing. In the paper, with expanded CEC (Chinese Event Corpus) as data source, prior knowledge of manifestation rules of event and relation as the guide, we propose an event extraction method based on knowledge-based rule of event manifestation, to achieve automatic building and improve text processing performance of event network.

1. Introduction

Text representation is an important issue in natural language processing, such as information retrieval and text classification. An appropriate representation not only can reflect text semantic, theme, and structure but also can improve the computational efficiency. In recent years, there is a tendency to use richer text representations than just keywords-based and concepts-based ones in the field of text information processing.

Event originated from cognitive science often appears in the literature of philosophy, cognitive science, linguistics, and artificial intelligence. It has been widely used in the computational linguistics as well as information retrieval and various NLP applications, which plays a special and important role in understanding text semantic. It not only contains specific correlations among a group of text elements but also indicates logical dependencies of things and attracts more and more attentions of researchers. Cognitive scientists believe that event is not only the basic unit of human cognizing and understanding objective world but also storage cell of proposition memory [1]. Most of the current natural language processing technologies lay particular stress on the theory of grammar structure, while ignoring the importance of semantic understanding, especially event semantic understanding [2]. Event-based text representation conforms to the rules of human cognition and natural language understanding.

Seen from present literature on event-based text representation we have consulted, there are the following main problems:

(1) The research on event-based text representation is still in its infancy; the thinking of event network is just beginning to blossom that it is necessary to be further explored.

(2) The operations and applications on event network need to be raised and further researched.

Against the shortcomings of current traditional text representation, the paper takes event as feature item of text and proposes an event-based text representation method. Event is regarded as semantic unit of text, and the events are connected by certain types of relations in the text, and these events imply correlations of linguistic units in the text by making the linguistic units (word, concept, sentence, etc.) as certain elements of event. It no longer regards text as an aggregation of independent words; consequently, the problem of "a bag of words" in classic text representation is

solved. Event network not only keeps semantic information of text and presents events and relations between events but also reflects importance and dynamic behavior of events. Compared to a traditional text representation, event network can express the higher granularity of semantic meaning, closer to the reality and easier for computers simulating text understanding and memorizing of human. It will provide new technology and method for semantic-based text information processing.

The paper is organized as follows: Section 2 introduces the related work. Section 3 constructs event network of Chinese texts in the field of emergencies. Section 4 evaluates the representing effect of event network. In Section 5, the formal definition of event network model for Chinese text is generalized by inducting and abstracting the instances of event network, and then the advantages of the model are analyzed. Finally, we summarize the paper and give an outlook of the future study.

2. Related Work

2.1. The Shortcoming of Traditional Text Representation. In the fields of information retrieval and natural language processing, the traditional text representation models mainly include the following: Boolean model [3], VSM (vector space model) [4, 5], BOW (bag of words) [6], LSI (latent semantic index) [7], LDA (latent Dirichlet allocation) [8], probability retrieval model [8], N-gram model [9], and language model [10].

Semantic information of text is composed of two parts: text component term (word, concept, sentence, etc.) and relationships between terms. Traditional text representation ignored the value of the order and relationships of the component terms on semantic expressing and assumed that the terms are independent, while, in fact, the semantic meaning of text is related to not only component terms and their frequency but also assembly rules and the order of terms, which means that the word-to-word and sentence-to-sentence relationships have an effect on text semantic. The same terms with same frequencies may express different semantic, such as the two following text snippets "Tom gave Mary a book as birthday gift" and "Mary gave Tom a book as birthday gift"; traditional text representation cannot express the difference between them [11]. Text representation based on word unit or concept unit will miss the information of relationships between terms, which will loss semantic meaning of text and result in failing to reflect higher level of semantic information. From the view of event semantic understanding, the above two text snippets express two different events.

In various texts, such as novel, opera, biography, and news reports, that contain many events, traditional text representation did not pay enough attention to event or represent event and relations appropriately. From the perspective of semantic understanding, linguists think that text is not only a group of attributes and concepts but also a describer of a series of events in a higher granularity; according to the thinking, these texts should be regarded as a group of events related by some relations, which is much closer to the laws of human cognition and understanding. From the perspective of formation of text, elementary language units (word, concept, sentence, etc.) form sentence by certain linguistic rules and sentences form a sequence of sentences or paragraph and then form text and express some semantic meaning and theme. Taking event as semantic unit of text and text component term as event-element only solves the problem of "a bag of words" but also expresses the higher level of semantic information.

2.2. Event-Oriented Text Representation. (Although the definitions of event are not unified in different applications, most of them emphasize two kinds of event attributes, action (verb or gerund) and characteristics of action (participant, location, time, etc.), so most researches are centered on verb and attributes of verb. In the paper, attribute of event is called event-element or element for short.) Looking from the current literature we have consulted, little research has been done on event-oriented text representation; the related work mainly includes the following.

Feng [12] proposed incident threading to represent English news reports at sentence level. The texts that describe the occurrence of a real-world happening are merged into a news incident, and incidents are organized in an incident threading by dependencies of predefined types. However, it does not do well in representing Chinese texts.

Glavaš and Šnajder [13] proposed an event-based text representation; however it only has temporal relation. Zhao-Man and Zong-Tian [14] proposed event lattice to represent narrative texts based on concept lattice. In the lattice, text is the object, event is the attribute, and binary relation is used to judge whether an event belongs to a text. Although lattice has precise mathematical properties, its describing power is weak, lacking the ability to express luxuriant relations. And obviously the event lattice has no meaning to one text. It is more suitable to represent inclusion between a group of texts and events than relations between events in a text.

Jian-fang and Yun-yu [15] expounded the thinking of event-based text representation in the paper named "The Research on Event-Oriented Text Representation." This paper discussed the feasibility and adaptability of event-based text representation for Chinese news reports at genre and arrangement of text. However it oversimplified relations between events, resulting in the fact that its representing power is weak. Thus there are still many issues need to be further studied.

Extracting events is the most important thing of event-based text representation. The three main approaches of extracting events are data-driven [16], knowledge-driven [17], and hybrid [18]. The accuracy is about 70 percent according to ACE (Automatic Content Extraction). The paper uses prior knowledge-guided approach.

3. Constructing Event Network of Chinese Text in the Field of Emergencies

Our experimental corpora, CEC (Chinese Event Corpus), are collected from Internet, the texts of which can be divided into five categories: earthquake, fire, traffic accidents, terrorist

Table 1: Statistics of annotated texts.

Text	300
Event	3977
Relation	2023
Coverage of text	85%

attack, and food poisoning according to the classification system of news report about emergency event [19]. Up to now, there are 500 texts in CEC, and 300 ones of them are human annotated event and relations. Some rules have been discovered based on the annotations using mining technology. KBR-EM (knowledge base of rule of event manifestation) has been constructed on CEC.

Verb plays an important role in semantic understanding; it is also core of event. As long as there is verb, it will involve maker and/or receiver of action, and certain regular collocation relationships will be established between action and involved entities; based on this, language would form various basic syntactic configurations and then explain construction of statement and relationship of vocabulary, and so forth. By annotating event on CEC, we find that event corresponds to verb or gerund, and 83% of these verbs or gerunds involve one or two entities, and arrangement of different text typology could affect the layout of events. The relations between events in text are as follows: some are contained in verb of sentence, some are expressed by conjunction (many conjunctions of text virtually show nontaxonomic relation between events, such as "because, therefore" indicates causation), and some are implied in the order of events (such as following relation); the experiment shows that two events will appear successively in text with great probability if there is a relation between them in reality. Our experiments show that events and relations meeting the above findings can cover 85% of entire text. Furthermore, following and causation are the largest number of relations, accounting for 81% of total relations. Statistics of the annotation are displayed in Table 1, where coverage of text is the ratio of event-contained sentences to total sentences. Thus it can be seen that event-based text representation will express text information appropriately.

The guidance of the KBR-EM modified the existing NLP tools (such as tokenizer, part-of-speech tagger, syntactic analyzer, and HowNet), and all of the programs are implemented in Java. Text is processed with word segmentation and POS tagging, syntax analysis and grammatical component tagging, identifying sentence and sentence components, and corresponding sentence or sentence components to event or event elements, regarding verb and gerund as trigger of event. and removing stop-using verbs, such as high-frequency verbs (be, do, have, etc.) and subjective verbs (feel, believe, etc.). Such events belong to stop-using events that are triggered by stop-using verbs; furthermore, stop-using events also include future events and negative events that are triggered by future-tense and negative-form verbs, respectively. Stop-using events should not be included in event network of text. Trigger-associated major components of action are other elements (time, place, subject/predicate-participant, etc.) of the event.

For the identified events, use electronic dictionary and ontology and make concept-climbing after mapping trigger of event into concept. Cluster event and generate event hierarchy by clustering based on the above climbed result, and taxonomic relations between events will be identified. According to the conjunction and other syntactic components of sentences where events are extracted from, consult the findings on relations mentioned above and identify nontaxonomic relations between events.

After identifying events and relations, event network is constructed as follows. Events in the text are arranged in a special directed graph. A named edge from event A to event B means that there is a relation between them in the text, either taxonomic (A is a B, forming multi-inheritance-allowed inheritance diagram) or nontaxonomic, such as causation (A leads to the happening of B), following (A precedes B in time), and composition (A is a part of B). And if there is more than one relation between event A and event B, then one relation is linked to one edge.

4. Experiment and Evaluating Representing Effect

Representing effect could measure whether a text representation method can represent information of original text appropriately and properly. The paper evaluates representing effect of event network with event recall rate (ER), event precision rate (EP), relation recall rate (RR), and relation precision rate (RP).

To compare between events and relations, the paper specifies some rules as follows:

(1) Two events are identical if and only if corresponding event elements are identical that are contained in the individual event.

(2) Two relations are identical if and only if corresponding items are identical that are contained in the individual relation tuple. For taxonomic relation $Is_{-a}(e_u, e_l)$, where e_u is superevent or upper-event, e_l is subevent or lower-event. For directed nontaxonomic relation $r\langle e_1, e_2\rangle$ and undirected nontaxonomic relation $r(e_1, e_2)$, where r is name of the relation, e_1 and e_2 are two events that are connected by the relation

Evaluating on event set of event networks of texts in the field of emergency, as shown in Figure 1, the average recall rate and precision rate are 82% and 88%, respectively. Evaluation of relation set is shown in Figure 2; the average recall rate and precision rate are 76% and 85%, respectively. Compared with previous method [15], the method constructs event network from tagged corpus with events, causation, and following relations, the resulting event network added another adjacent relation and event-element-shared relation. Its event recall and precision rate will be higher, and events contained in the event network can be viewed as complete and correct in theory. According to the findings described in Section 3, nontaxonomic relation recall rate should be at least 81%. However, there are large amount of redundancy

Table 2: Comparison for event network and incident threading.

	ER	EP	Event F_measure	RR	RP	Relation F_measure
Event network	82%	88%	84%	76%	85%	80%
Incident threading	41%	98%	58%	18%	92%	30%

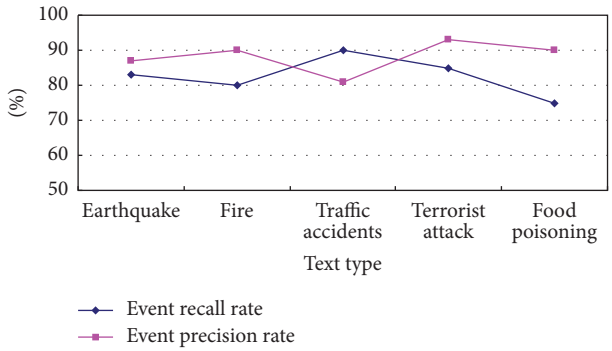


Figure 1: Evaluation of event set.

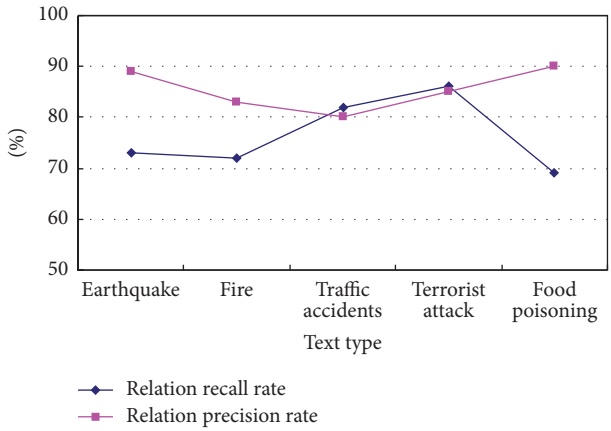


Figure 2: Evaluation of relation set.

and error in adjacent and event-element-shared relation; for example, adjacent relation could actually be following or with no meaningful relation, and event-element-shared relation is too general to specify a relation. So relation precision rate of this method is far inferior to the paper.

Incident threading [12] did well in representing preprocessed grouped English news texts; however, it is less suitable for Chinese text than event network. Evaluating the two representation methods is shown in Table 2.

5. Event Network Model for Chinese Text

An event network contains one or more events that are connected by relations. Events in the network are arranged in a graph, and two events are directly connected by one or more directed/undirected edges (the number of edges depends on the number of relations between the two events) and have some relations. The text representation method is called event network. Though constructing event networks of a large number of texts, we discover that event network is different from general directed digraph. There is information on its each node and each edge, and multiple edges may exist between two nodes. The formal event network model is defined as following by generalizing and abstracting instances of event networks.

Definition 1 (event network). The tuple EN $= (E, R \diamond (R_T, R_{\mathrm{NT}}))$ is called event network that meets the following conditions:

(1) $E = \{e\}$ is nonempty node set, called event set.

(2) $R \diamond (R_T, R_{\mathrm{NT}})$ is edge set, called relation set.

R includes taxonomic relation R_T and nontaxonomic relation R_{NT}. Taxonomic relation $R_T = \{Is_{-a}(e_u, e_l) \mid e_u \in E, e_l \in E\}$ forms multi-inheritance-allowed inheritance diagram, where e_u is superevent and e_l is subevent. R_{NT} forms special graph structure, including directed $R_{\mathrm{NT}} = \{r\langle e_1, e_2\rangle \mid e_u \in E, e_l \in E\}$ and undirected $R_{\mathrm{NT}} = \{r(e_1, e_2) \mid e_u \in E, e_l \in E\}$, where the relation between event e_1 and e_2 is named as r.

Event network can be seen as directed graph. It not only keeps semantic information of text and represents events and relations between events but also reflects importance, dynamic behavior, and state changing of events. Compared with traditional text representation such as VSM, the salient advantage of event network is that it implies correlations among linguistic units of the text in its events, which not only solves the problem of "a bag of words" but also inflects the higher granularity of semantic meaning. Meanwhile relations link events together that can express logical dependencies of things and reflect the occurrence and development process of event.

Event network is a directed graph with information on its nodes and edges. Using all information, various calculations can be done on it by considering some properties of directed graph; for example, an event network can be clustered according to the similarity of events, partitioned into hierarchical structure with different threshold value, and reduced according to importance of event or can keep some other properties. The similarity of texts can be calculated according to the matching of their individual event network; some knowledge can be obtained through mining frequent and simultaneous event elements in multiple event networks. These calculations must meet not only properties of graph but also meaning of information on nodes and edges of event network, so the unique properties and special computation model of event network need to be researched. Establishing abstract operations on event network, some problems will be solved by mathematical methods, which are a kind of good

form for semantic calculation and will support event-based text information processing.

6. Conclusions and the Future Work

The paper introduced requirement of event-based text representation. The formal event network model for Chinese text is defined by abstracting instances of event network on CEC texts. The difference between event network and traditional text representation is that event network keeps semantic information of text, no longer regards text as an aggregation of independent words, and solves the problem of "a bag of words." In addition, it reflects relations among events, importance, and dynamic behavior of event. Our experiments demonstrate the feasibility, adaptability, and advantage of event network as a text representation method.

In the future work, we will study computations on event network by using graph theory, clustering, formal concept analysis, granular computing, and so forth, considering particularity of the model. In this way, various applications of text will be solved by mathematical methods. Theoretical model and method support for present text information processing based on semantic meaning will be provided.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work is supported by Science and Technology Innovation Project of Chinese Ministry of Culture (2015KJCXXM19) and Foundation for University Youth Teachers of Shanghai (ZZGCD15002).

References

[1] P. Yun-he and W. Geng, "An introduction to intelligent-computing oriented memory theory," *Journal of Computer Research and Development*, vol. 31, pp. 37–42, 1999.

[2] L. Zhong, *The theory of BSCM presented and the implementing of the natural language understanding studied*, East China Normal University, 2004.

[3] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, Addison-Wesley-Longman, 1st edition, 1999.

[4] G. Salton, A. Wong, and C. S. Yang, "A Vector Space Model for Automatic Indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975.

[5] D. D. Lewis, "Evaluation of phrasal and clustered representations on a text categorization task [A]," in *Proceedings of the Fifteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 37–50, 1992.

[6] Bag-of-words model [DB/OL], http://en.wikipedia.org/wiki/Bag_of_words_model.

[7] T. K. Landauer and M. L. Littman, "Fully automatic crosslanguage document retrieval using latent semantic indexing," in *Proceedings of the Sixth Annual Conference of the UW Centre for the New Oxford English Dictionary and Text Research*, pp. 31–38, Waterloo Ontario, 1990, http://www.es.duke.edu/~mlittman/docs/x-lang.ps.

[8] M. David and Y. Blei Andrew, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[9] T. Chew Lim, S. Sung Yuan, Y. Zhaohui et al., *Text Retrieval from Document Images Based on N-Gram Algorithm*, http://citeseer.nj.nec.com/400555.html.

[10] J. Ponte and W. Croft, "Language Modeling Approach to Information Retrieval," in *Proceedings of SIGIR1998'*, pp. 275–281.

[11] J.-F. Shan, Z.-T. Liu, J.-F. Fu, and Z.-M. Zhong, "Important event extraction of Chinese document based on small world model," in *Proceedings of 2009 IEEE International Conference on Intelligent Computing and Intelligent Systems, ICIS 2009*, pp. 146–150, chn, November 2009.

[12] A. Feng, *Events Threading*, University of Massachusetts, 2008.

[13] G. Glavaš and J. Šnajder, "Event graphs for information retrieval and multi-document summarization," *Expert Systems with Applications*, vol. 41, no. 15, pp. 6904–6916, 2014.

[14] Z. Zhao-Man and L. Zong-Tian, "Events-based Text Similarity Computing," *Journal of Guangxi Normal University (Natural Science Edition)*, vol. 27, no. 1, pp. 149–152, 2009.

[15] Sh. Jian-fang and Sh. Yun-yu, "Research on event network for Chinese text," in *Proceeding of the Interna-tional Symposium on Information Technology Convergence, ISITC2016*, pp. 473–482, Shanghai, China, 2016.

[16] M. Okamoto and M. Kikuchi, "Discovering volatile events in your neighborhood: Local-area topic extraction from blog entries," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5839, pp. 181–192, 2009.

[17] E. Minkov, "Event extraction using structured learning and rich domain knowledge: Application across domains and data sources," *ACM Transactions on Intelligent Systems and Technology*, vol. 7, no. 2, article no. 16, 2015.

[18] S. Kuptabut and P. Netisopakul, "Event extraction using ontology directed semantic grammar," *Journal of Information Science and Engineering*, vol. 32, no. 1, pp. 79–96, 2016.

[19] Y. Li-ying, L. Hong-juan, and Z. Yong-kui, "The research on classification system of emergency news corpus," in *Proceeding of the monograph of the 25th academic annual conference of Chinese Informatics Association, frontier of Chinese information processing*, Tsinghua University Press, Beijing, 2006.

Hindawi
International Journal of Digital Multimedia Broadcasting
Volume 2017, Article ID 1386461, 7 pages
https://doi.org/10.1155/2017/1386461

Research Article

A Novel Preferential Diffusion Recommendation Algorithm Based on User's Nearest Neighbors

Fuguo Zhang,[1,2] Yehuan Liu,[1] and Qinqiao Xiong[1]

[1] *School of Information Technology, Jiangxi University of Finance & Economics, Nanchang 330013, China*
[2] *Research Institution for Information Resource Management, Jiangxi University of Finance & Economics, Nanchang 330013, China*

Correspondence should be addressed to Fuguo Zhang; redbird_mail@163.com

Received 1 March 2017; Accepted 11 April 2017; Published 4 May 2017

Academic Editor: Hyo-Jong Lee

Recommender system is a very efficient way to deal with the problem of information overload for online users. In recent years, network based recommendation algorithms have demonstrated much better performance than the standard collaborative filtering methods. However, most of network based algorithms do not give a high enough weight to the influence of the target user's nearest neighbors in the resource diffusion process, while a user or an object with high degree will obtain larger influence in the standard mass diffusion algorithm. In this paper, we propose a novel preferential diffusion recommendation algorithm considering the significance of the target user's nearest neighbors and evaluate it in the three real-world data sets: MovieLens 100k, MovieLens 1M, and Epinions. Experiments results demonstrate that the novel preferential diffusion recommendation algorithm based on user's nearest neighbors can significantly improve the recommendation accuracy and diversity.

1. Introduction

With the rapid development of Internet in the past years, the amount of online information increases at an exponential speed, which leads to information overload problem. When faced with vast amount of information, we can hardly find the valuable information accurately and quickly. The personalized recommender system is one of the most effective tools to resolve this problem, and it also can help enterprises make the users' potential demand a realistic demand [1, 2].

To date, various recommendation methods have been proposed and developed. One of the most successful recommender system methods is based on the collaborative filtering technique [3–5]. Recently, some physical methods, such as mass diffusion [6–9] and heat conduction [10, 11], have found applications in personalized recommendation. Standard mass diffusion algorithm applied the three-step mass diffusion starting from the target user on a user-object bipartite network, which accurately outperforms the standard collaborative filtering methods [1]. Many different bipartite network based methods [12] are proposed to achieve even better recommendation performance. In [6], Zhou et al. proposed a hybrid method by combining the mass diffusion and heat conduction to solve the apparent diversity-accuracy dilemma of recommender systems. Motivated by enhancing the preferential diffusion algorithm's ability to find unpopular and niche objects, the preferential diffusion has been designed in [9]. Moreover, Zhang and Zeng proposed a strategy to adding some virtual connections to the networks, which is useful to deal with the cold start problem in recommender system [13].

However, all these methods do not give a high enough weight to the influence of the target user's nearest neighbors in the resource diffusion process. As we all know, birds of a feather flock together. The user's nearest neighbors are the ones who have similar taste with the given user. Therefore we introduce a novel preferential diffusion recommendation algorithm considering the significance of the target user's nearest neighbors in the diffusion process.

2. Methods

A recommender system can be represented by a bipartite network $G(U, O, E)$, where $U = \{u_1, u_2, \ldots, u_m\}$, $O = \{o_1, o_2, \ldots, o_n\}$, and $E = \{e_1, e_2, \ldots, e_q\}$ are the sets of users, objects,

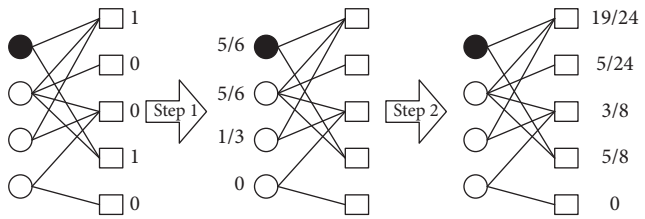


Figure 1: Standard mass diffusion algorithm at work on the bipartite user-object network. Users are shown as circles and objects are squares. The target user is indicated by the black circle.

and links, respectively [7]. Denote by $A_{m\times n}$ the adjacency matrix, where the element $a_{ia} = 1$ if the user i has selected the object a and $a_{ia} = 0$ otherwise.

2.1. Standard Mass Diffusion Recommendation Algorithm. As is shown in Figure 1, the standard mass diffusion (SMD) algorithm is equivalent to a three-step random walk process. At first, objects in the bipartite network are assigned an initial resource f, with $f^i = \{f_1^i, f_2^i, \ldots, f_\alpha^i, \ldots, f_n^i\}$ for the target user i. For simplicity, if an object is collected by the user i, its initial resource is assigned to be 1, otherwise it is assigned to be 0. That is to say, the initial resource vector f can be written as

$$f_\alpha^i = a_{i\alpha}. \tag{1}$$

Then, each object's resource is redistributed to the user who has collected the object averagely, and the user's resource is the sum of the resources received from objects. At last, each user's resource was reallocated to the objects which he has collected averagely. The final score of the object's resource can be calculated via the transformation $f' = Wf$, where W is the resource transfer matrix.

$$w_{\alpha\beta} = \frac{1}{k_\beta} \sum_{l=1}^{m} \frac{a_{l\alpha} a_{l\beta}}{k_l}, \tag{2}$$

where k_β is the degree of the object β and k_l is the degree of the user l.

2.2. The Novel Preferential Diffusion Algorithm Based on User's Nearest Neighbors. Following on from previous research [14], the diffusion process of the novel preferential diffusion recommendation algorithm based on user's nearest neighbors (NNMD) is shown in Figure 2. At first, we calculate the Jaccard similarities between the target user i and the other users to get the top N similar neighbors. The formula of Jaccard similarity reads

$$J_{ij} = \frac{\left|N_i \cap N_j\right|}{\left|N_i \cup N_j\right|}, \tag{3}$$

where J_{ij} is the Jaccard similarity between user i and user j and N_i and N_j are the user neighbors set of user i and user j, respectively. Then we can get the objects' initial resource denoted by the vector f_1, with $f_1^i = \{f_{11}^i, f_{12}^i, \ldots, f_{1\alpha}^i, \ldots, f_{1n}^i\}$ for the target user i. f_1 can be written as

$$f_{1\alpha}^i = a_{i\alpha} + \sum_{k\in U} a_{k\alpha}, \tag{4}$$

where U is the nearest neighbors set of the target user i. In Figure 2, $f_1 = (3, 1, 2, 2, 0)$. But only the objects which the target user i has selected can distribute the resources to users and then redistribute them via the transformation

$$f_2 = W\left(f_1 f\right), \tag{5}$$

where W is the same as (2). In Figure 2, $f_2 = (2, 0.5, 1, 1.5, 0)$. Finally, we use the linear combination the resources vectors f_1 and f_2 to get the last objects' resources vector F. That is to say,

$$F = \alpha f_1 + (1 - \alpha) f_2, \tag{6}$$

where α is a variable parameter from 0 to 1.

3. Data and Metrics

3.1. Data. To test the algorithmic performance, we use three benchmark data sets as shown in Table 1. The sparsity of these data sets is shown in the last column of Table 1. They are very sparse, especially Epinions data set. MovieLens 100k and MovieLens 1M data sets [15] were collected by the GroupLens research group. They consist of 100000 ratings from 943 users on 1682 different movies and 1000209 ratings from 6040 users on 3952 different movies, respectively. The ratings are integer numbers in the range of 1 to 5 scales. The Epinions data set [16] consists of 22166 users, 296277 objects, and 922267 ratings. It is noted that Epinions data set is highly sparse. Users only rate a small number of items in the system, and, in order to get better results, we delete those users and objects with degree less than 7. Finally, we get a new data set which consists of 4066 users, 7649 objects, and 154122 ratings. We randomly divide the data sets into two parts: the training set E^T contains 80% of the data and the remaining 20% of data constitutes the probe set E^P.

3.2. Metrics. There has been considerable research in the area of recommender systems evaluation. Accuracy is the most important aspect in evaluating the recommendation algorithmic performance. In this paper, we use ranking score [8] to measure the ability of a recommendation algorithm to generate a ranking list of the target user's uncollected objects that matches the users' preference. For the target user u_i, the recommendation algorithm will return u_i a ranking list of all his unselected objects and, according to E^P, if u_i has selected the object o_j and o_j is at r_{ij}th place in the ranking list, we say the position of o_j is

$$R_{ij} = \frac{r_{ij}}{L}, \tag{7}$$

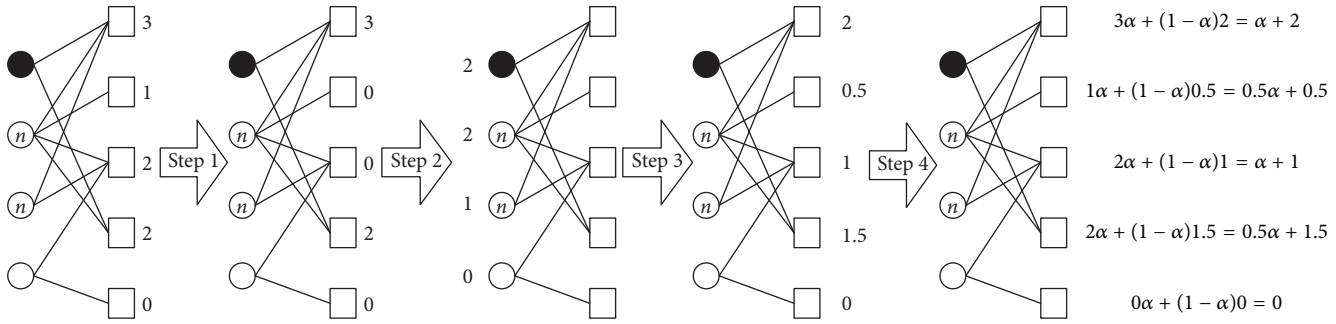


Figure 2: A novel preferential diffusion algorithm by user's nearest neighbors at work on the bipartite user-object network. Users are shown as circles and objects are squares. The target user is indicated by the black circle and the nearest neighbors of the target user are the circles which have a letter "n" in them and α is the variable parameter.

Table 1: Basic properties of the three data sets and the sparsity is defined as $E/(N_u N_o)$.

Network	E	N_u	N_o	Sparsity
MovieLens 100k	100000	943	1682	0.063
MovieLens 1M	1000209	6040	3952	0.042
Epinions	154122	4066	7649	0.005

Table 2: The four cases of the unselected objects of the target user in the recommender system.

User likes	Recommender system recommended	Recommender system did not recommend
Likes	C_{tp}	C_{fn}
Does not like	C_{fp}	C_{tn}

where L is the number of his unselected objects. We obtain the mean value of all the user-object ranking scores in E^P; namely,

$$R = \frac{1}{|E^P|} \sum_{ij \in E^P} R_{ij}. \tag{8}$$

Clearly, the larger the ranking score, the lower the algorithm's accuracy and vice versa.

In the practical recommender system, we may consider the number of objects that users like in the recommendation list. Therefore, we take another accuracy metric called precision. For a target object o_j and user u_i, there are four cases in the recommender system. The first is that the recommender system recommended the object and user likes it. The second is that recommender system recommended the object but the user does not like it. The third is that the user likes the object but the recommender system did not recommend it. Finally is the case that the user does not like the object and the recommender system did not recommend it. As is shown in Table 2, C_{tp}, C_{fn}, C_{fp}, and C_{tn} denote the number of the objects in the four cases.

For a target user u_i, the precision of recommendation $P_i(L)$ is defined as

$$P_i(L) = \frac{C_{tp}}{L} = \frac{C_{tp}}{C_{tp} + C_{fp}}. \tag{9}$$

We obtain the mean precision $P(L)$ of all the users in the recommender system. Besides accuracy, diversity is taken into account as another important aspect to evaluate the recommendation algorithm. There are two kinds of diversity. One is called intrauser-diversity [17]; the other is called interuser-diversity [18]. In this paper, we consider the interuser-diversity. It considers the different objects between users in the recommendation list. For two users u_i and u_j, the differences can use be measured by the Hamming distance [18]:

$$H_{ij}(L) = 1 - \frac{S_{ij}(L)}{L}, \tag{10}$$

where $S_{ij}(L)$ is the number of common objects between u_i and u_j in the recommendation list and L is the length of the recommendation list. Clearly, if u_i and u_j have the same recommendation list, $H_{ij}(L) = 0$, while if the recommendation lists are completely different, $H_{ij}(L) = 1$.

In reality, it has been found that a recommender system which has a high accuracy might not be satisfied by the users [19]. For example, for a film website, recommending the popular films to the users may not always be the best recommendation, because users might have already seen those films in other ways. A good recommender system can find the objects that match the users' preferences and are unlikely to be already known. As a result, the novelty is also often used in evaluating the recommendation algorithmic performance.

The average degree of objects in the recommendation list is widely used to identify the novelty of a recommender system [20], which is defined by

$$N(L) = \frac{1}{ML} \sum_{u} \sum_{o_\partial \in O_R^i} k_{o_\partial}, \tag{11}$$

where M is the number of users, O_R^i is the recommendation list for user u_i, and k_{o_∂} is the degree of the object o_∂.

Table 3: Algorithmic performance for MovieLens 100k, MovieLens 1M, and Epinions data. The precision, interuser-diversity, and novelty are corresponding to $L = 20$. The parameters for NNMD are $N = 50$ and $\alpha = 0.9$ in MovieLens 100k and MovieLens 1M, while, in Epinions, the parameters are $N = 10$ and $\alpha = 0$. The entries corresponding to the best performance over all methods are emphasized in bold.

Data set	Algorithms	Ranking score	Precision	Interuser-diversity	Novelty
MovieLens 100k	NNMD	**0.059537**	**0.2242**	**0.8401**	**237**
	SMD	0.069011	0.1971	0.6970	279
MovieLens 1M	NNMD	**0.077039**	**0.2726**	**0.8816**	**1340**
	SMD	0.095269	0.1949	0.5865	1828
Epinions	NNMD	**0.180439**	**0.0374**	**0.6787**	**204**
	SMD	0.181141	0.0357	0.6743	205

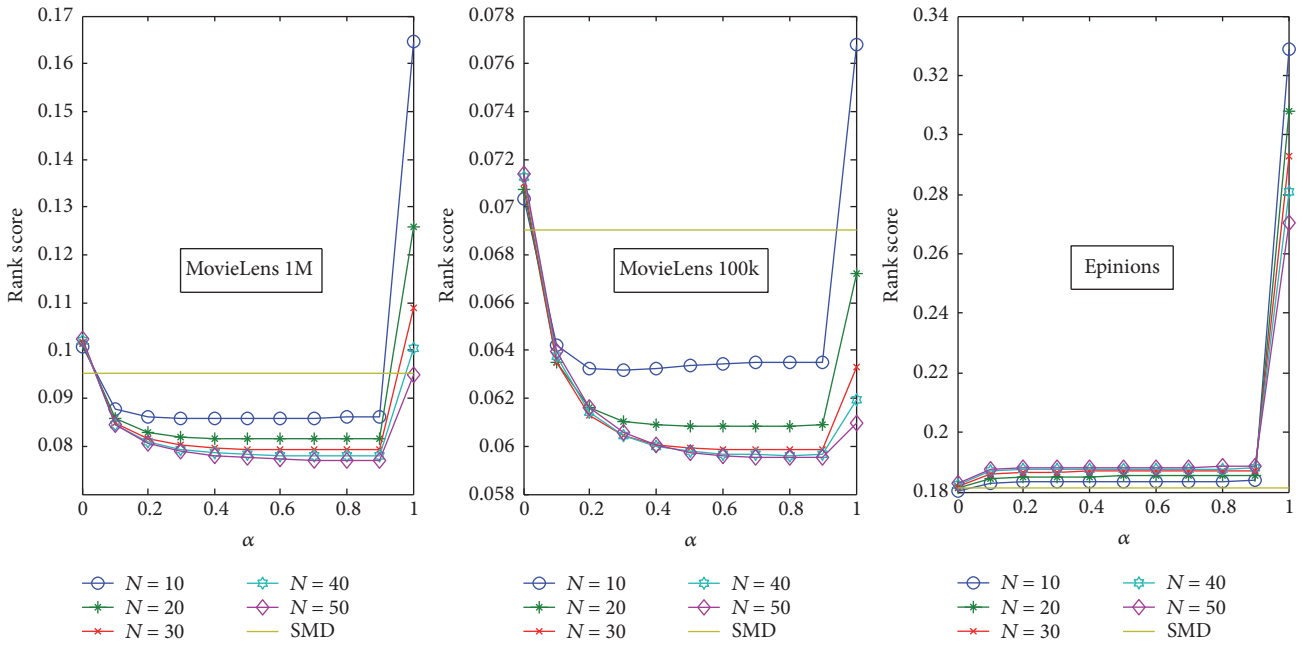


Figure 3: The overall ranking score of the NNMD algorithm under different N and α in MovieLens 100k, MovieLens 1M, and Epinions data set and the ranking score of the SMD algorithm in MovieLens 100k, MovieLens 1M, and Epinions.

4. Results and Discussion

In our first set of experiments, we compare the ranking score of the NNMD algorithm under different α and top N (N is the number of the target user's nearest neighbors) with that of the SMD algorithm. The results on MovieLens 100k, MovieLens 1M, and Epinions data are reported in Figure 3. Clearly, we can see that in MovieLens 100k and MovieLens 1M, with the increase of N, the rank score is smaller and smaller; that is to say, the recommendation accuracy is getting better and better. However, when N is more than 30, the change of rank score is very small. Moreover, as long as α is not equal to 0 or 1, the rank score of our method is better than that of the SMD algorithm. It is interesting to note that the optimal parameters of our method are the same in MovieLens 100k and MovieLens 1M, which are $N = 50$ and $\alpha = 0.9$, while, in Epinions, the improvement of the rank score is not significant. When α is greater than 0 or N is greater than 20 the rank score of the NNMD algorithm is a little worse than that of the SMD algorithm, and, with the change of α and N, the rank scores of the two algorithms are almost the same. But when N is less than 20 and $\alpha = 0$, the rank score of our method is getting better than that of the SMD algorithm. Clearly, we can get the optimal parameters $N = 10$ and $\alpha = 0$ in Epinions.

Then we examined the performance in precision, interuser-diversity, and novelty of our novel algorithm at the optimal parameters N and α. Summaries of the results for all algorithms and metrics on MovieLens 100k, MovieLens 1M, and Epinions data sets are shown in Table 3. The optimal parameters are subject to the lowest ranking score. The other three metrics, namely, precision, interuser-diversity, and novelty, are obtained at the optimal parameters. Clearly, the NNMD algorithm outperforms the SMD algorithm over all four evaluation metrics.

The comparison of precision between NNMD and SMD in three data sets under different length of recommendation list is shown in Figure 4. It clearly indicates that the precision

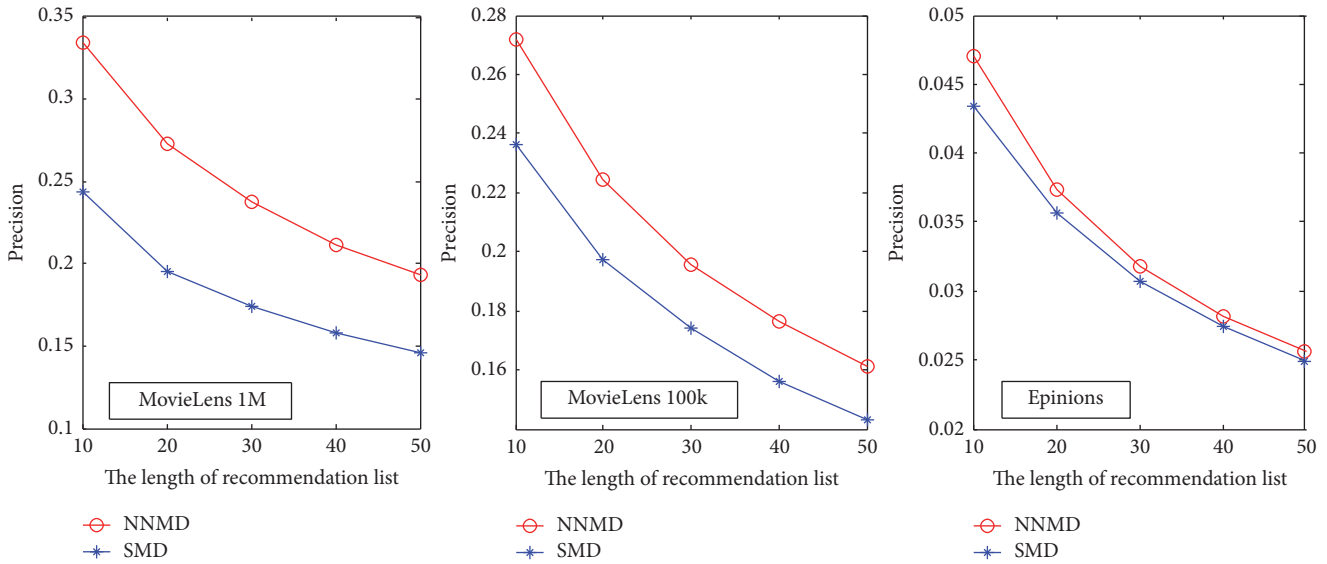


FIGURE 4: The precision of NNMD and SMD algorithm in MovieLens 100k, MovieLens 1M, and Epinions under different length of recommendation list. The parameters for the NNMD algorithm are $N = 50$ and $\alpha = 0.9$ in MovieLens 100k and MovieLens 1M while they are $N = 10$ and $\alpha = 0$ in Epinions.

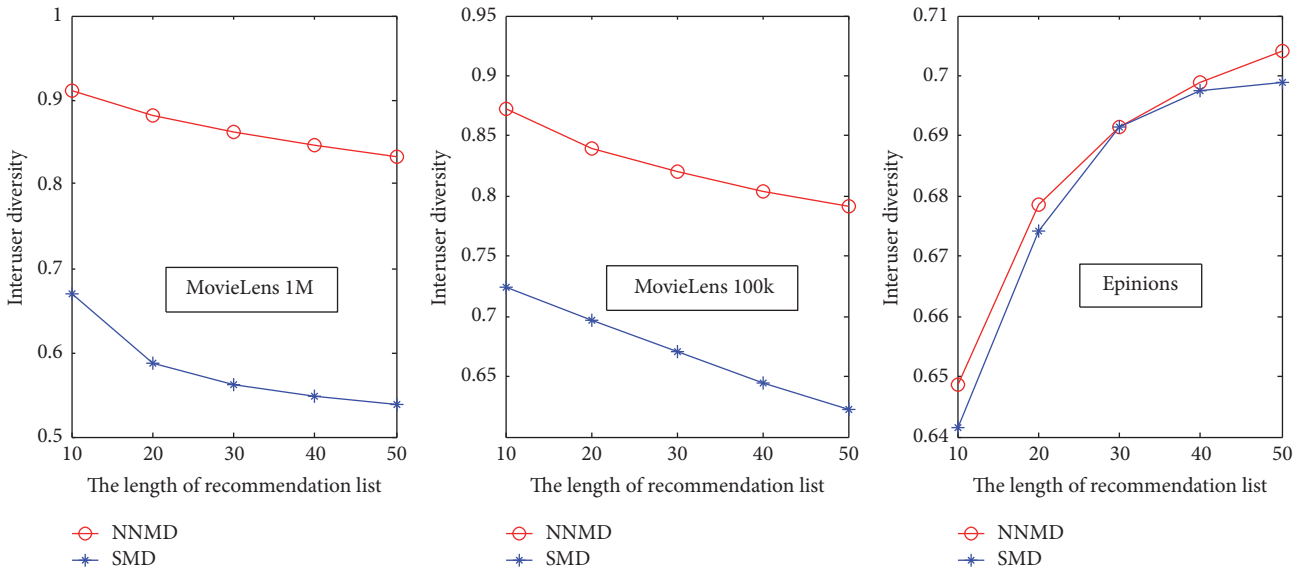


FIGURE 5: The interuser-diversity of the NNMD and SMD algorithm in MovieLens 100k, MovieLens 1M, and Epinions under different length of recommendation list. The parameters for the NNMD algorithm are $N = 50$ and $\alpha = 0.9$ in MovieLens 100k and MovieLens 1M while they are $N = 10$ and $\alpha = 0$ in Epinions.

of the NNMD algorithm is better than that of the NMD algorithm in all the three data sets and it has a very significant improvement in MovieLens 100k and MovieLens 1M. That is to say, our method can recommend objects for users more accurately.

Figure 5 shows the comparison of interuser-diversity between our method NNMD and SMD in three data sets under different length of recommendation list. It clearly shows that interuser-diversity of our NNMD algorithm is better than that of the SMD algorithm in all the three data sets, especially in MovieLens 100k and MovieLens 1M. In other words, the objects in the recommendation list of our method are more different between users.

Figure 6 shows the comparison of novelty between our method NNMD and SMD in three data sets under different length of recommendation list. It clearly indicates that the novelty of our method is much better than the SMD in MovieLens 100k and MovieLens 1M, while, in Epinions, the results of the two algorithms are very similar, but our method also has a little improvement than that of the SMD algorithm.

In summary, the recommendation performance of our method is better than that of the standard mass diffusion. In

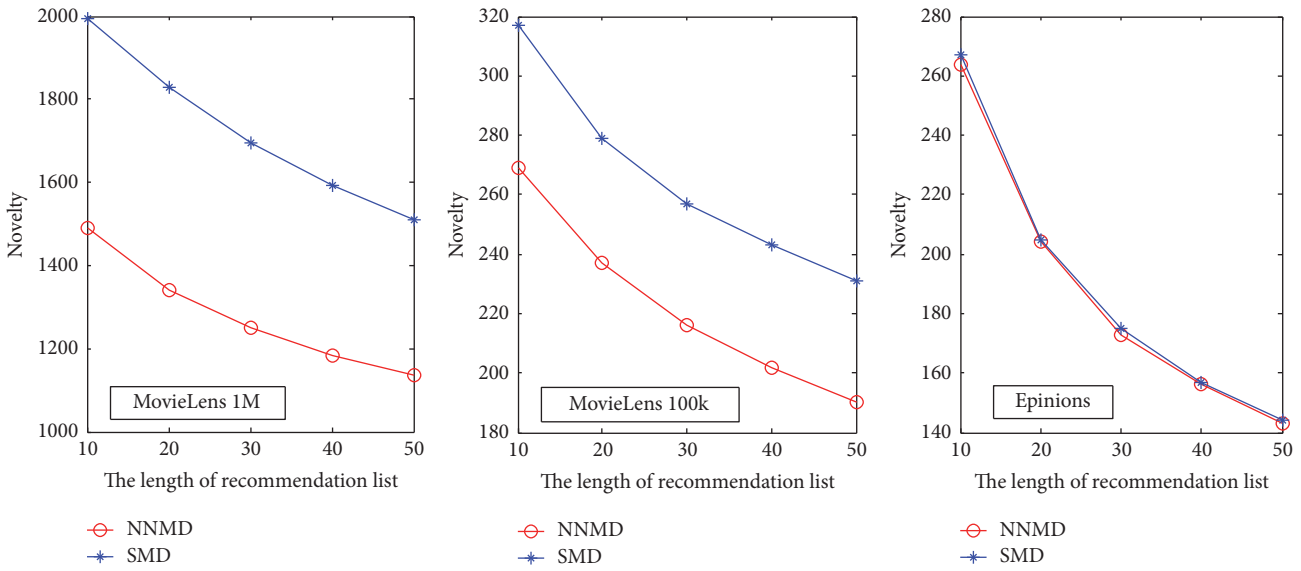


Figure 6: The novelty of NNMD and SMD algorithm in MovieLens 100k, MovieLens 1M, and Epinions under different length of recommendation list. The parameters for the NNMD algorithm are $N = 50$ and $\alpha = 0.9$ in MovieLens 100k and MovieLens 1M, while they are $N = 10$ and $\alpha = 0$ in Epinions.

particular, the precision of our method increases an average of 13.27% percent compared to that of the SMD in MovieLens 100k and increases an average of 35.9% percent in MovieLens 1M and increases an average of 4.47% percent in Epinions. Although the improvement of the algorithmic performance in some aspects is not significant in Epinions data set, the reason may be that the data is so sparse that the novel algorithm cannot get the proper user's nearest neighbors and it affects our algorithmic performance.

5. Conclusion and Future Work

Most of network based recommendation algorithms have a tendency to recommend popular objects to the users [1] because the object with high degree has a significant influence in the resource diffusion process. In this paper we propose a novel preferential diffusion recommendation algorithm based on user's nearest neighbors which give a high weight to the influence of the target user's nearest neighbors in the resource diffusion process. Experimental results based on MovieLens 100k, MovieLens 1M, and Epinions data set show that making a suitable adjustment in the parameter α or the size of the user's nearest neighbors set can help recommendation algorithm get a better recommendation performance. It can not only provide more accurate recommendations but also generate more diverse and novel recommendations.

For future work, we intend to consider the level of rating between user and his nearest neighbors. Moreover, we will use the trust data [21, 22] in the network, because it can be used to find the nearest neighbors more accurately in high sparse data set, and it may have a better recommendation performance.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is partially supported by National Natural Science Foundation of China (Grant nos. 71361012 and 71363022), by National Science Foundation of Jiangxi, China (no. 20161BAB201029), and by the Foundation of Jiangxi Provincial Department of Education (no. GJJ. 150446).

References

[1] T. Zhoua, Z. Kuscsik, J. Liu, M. Medo, J. R. Wakeling, and Y. Zhang, "Solving the apparent diversity-accuracy dilemma of recommender systems," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 10, pp. 4511–4515, 2010.

[2] L. Lü, M. Medo, C. H. Yeung, Y. Zhang, Z. Zhang, and T. Zhou, "Recommender systems," *Physics Reports*, vol. 519, no. 1, pp. 1–49, 2012.

[3] J. A. Konstan, B. N. Miller, D. Maltz, J. L. Herlocker, L. R. Gordon, and J. Riedl, "Applying collaborative filtering to usenet news," *Communications of the ACM*, vol. 40, no. 3, pp. 77–87, 1997.

[4] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," in *Proceedings of the 14th conference on Uncertainty in Artificial Intelligence*, pp. 43–52, 1998.

[5] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, 2005.

[6] T. Zhou, J. Ren, M. Medo, and Y. Zhang, "Bipartite network projection and personal recommendation," *Physical Review E*, vol. 76, no. 4, Article ID e046115, 2007.

[7] M. S. Shang, L. Lu, Y. C. Zhang, and T. Zhou, "Empirical analysis of web-based user-object bipartite networks," *Europhysics Letters*, vol. 90, no. 4, Article ID e48006, 2010.

[8] A. Zeng, A. Vidmer, M. Medo, and Y.-C. Zhang, "Information filtering by similarity-preferential diffusion processes," *EPL*, vol. 105, no. 5, Article ID e58002, 2014.

[9] L. Lu and W. Liu, "Information filtering via preferential diffusion," *Physical Review E*, vol. 83, no. 6, Article ID e066119, 2011.

[10] Y. C. Zhang, M. Blattner, and Y. K. Yu, "Heat conduction process On community networks as a recommendation model," *Physical Review Letters*, vol. 99, no. 10, Article ID 154301, 2007.

[11] J. G. Liu, T. Zhou, and Q. Guo, "Information filtering via biased heat conduction," *Physical Review E*, vol. 84, no. 3, Article ID e037101, 2011.

[12] F. Yu, A. Zeng, S. Gillard, and M. Medo, "Network-based recommendation algorithms: a review," *Physica A: Statistical Mechanics and Its Applications*, vol. 452, pp. 192–208, 2016.

[13] F. Zhang and A. Zeng, "Improving information filtering via network manipulation," *Europhysics Letters*, vol. 100, no. 5, Article ID 58005, 2012.

[14] F. G. Zhang, Y. H. Liu, and Q. Q. Xiong, "A novel mass diffusion recommendation algorithm based on user's nearest neighbors," in *Proceedings of the International Symposium on Information Technology Convergence*, 2016.

[15] http://www.grouplens.org/.

[16] http://www.epinions.com/.

[17] T. Zhou, R. Q. Su, R. R. Liu, L. L. Jiang, B. H. Wang, and Y. Zhang, "Accurate and diverse recommendations via eliminating redundant correlations," *New Journal of Physics*, vol. 11, Article ID 123008, 2009.

[18] T. Zhou, L. L. Jiang, R. Q. Su, and Y. C. Zhang, "Effect of initial configuration on network-based recommendation," *Europhysics Letters*, vol. 81, no. 5, pp. 58004–58007, 2008.

[19] M. Ge, C. Delgado-Battenfeld, and D. Jannach, "Beyond accuracy: evaluating recommender systems by coverage and serendipity," in *Proceedings of the 4th ACM Conference on Recommender Systems (RecSys '10)*, pp. 257–260, Barcelona, Spain, September 2010.

[20] Z.-K. Zhang, C. Liu, Y.-C. Zhang, and T. Zhou, "Solving the cold-start problem in recommender systems with social tags," *Europhysics Letters*, vol. 92, no. 2, Article ID 28002, 2010.

[21] C. Martinez-Cruz, C. Porcel, J. Bernabé-Moreno, and E. Herrera-Viedma, "A model to represent users trust in recommender systems using ontologies and fuzzy linguistic modeling," *Information Sciences*, vol. 311, pp. 102–118, 2015.

[22] X. Qian, H. Feng, G. Zhao, and T. Mei, "Personalized recommendation combining user interest and social circle," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 7, pp. 1763–1777, 2014.