Augmented reality based real-time subcutaneous vein imaging system

Danni Ai,[1] **Jian Yang,**[1,*] **Jingfan Fan,**[1] **Yitian Zhao,**[1] **Xianzheng Song,**[1] **Jianbing Shen,**[2] **Ling Shao,**[3] **and Yongtian Wang**[1]

[1]*Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Electronics, Beijing Institute of Technology, Beijing 100081, China*
[2]*Beijing Laboratory of Intelligent Information Technology, School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China*
[3]*Department of Computer Science and Digital Technologies, Northumbria University, Newcastle upon Tyne NE1 8ST, U.K.*

[*]jyang@bit.edu.cn

Abstract: A novel 3D reconstruction and fast imaging system for subcutaneous veins by augmented reality is presented. The study was performed to reduce the failure rate and time required in intravenous injection by providing augmented vein structures that back-project superimposed veins on the skin surface of the hand. Images of the subcutaneous vein are captured by two industrial cameras with extra reflective near-infrared lights. The veins are then segmented by a multiple-feature clustering method. Vein structures captured by the two cameras are matched and reconstructed based on the epipolar constraint and homographic property. The skin surface is reconstructed by active structured light with spatial encoding values and fusion displayed with the reconstructed vein. The vein and skin surface are both reconstructed in the 3D space. Results show that the structures can be precisely back-projected to the back of the hand for further augmented display and visualization. The overall system performance is evaluated in terms of vein segmentation, accuracy of vein matching, feature points distance error, duration times, accuracy of skin reconstruction, and augmented display. All experiments are validated with sets of real vein data. The imaging and augmented system produces good imaging and augmented reality results with high speed.

© 2016 Optical Society of America

References and links

1. P. Deepa, K. Mohanavelu, B.S. Sundersheshu, V.C. Padaki, "Vein Identification and Localization for Automated Intravenous Drug Delivery System," Volume 292 of the series Communications in Computer and Information Science, (2012), pp. 270–281.
2. N.J. Cuper, J.H. Klaessens, J.E. Jaspers, R. de Roode, H.J. Noordmans, J.C. de Graaff, R.M. Verdaasdonk, "The use of near-infrared light for safe and effective visualization of subsurface blood vessels to facilitate blood withdrawal in children," Med. Eng. Phys **35**(4), 433–440 (2013).
3. D. Delpy, "Quantification in tissue near-infrared spectroscopy," Philos Trans R Soc Lond B. **352**(1354), 649–659 (1997).
4. H.D. Zeman, G. Lovhoiden and C. Vrancken, "Prototype vein contrast enhancer," Proc. SPIE **5318**, 39–49 (2004).
5. https://www.christiemed.com/products/veinviewer-models/veinviewer-vision

6. AccuVein AV400: The only handheld, non-contact vein illumination device, http://medidyne.dk/wp-content/uploads/ENG-accuvein-data-sheet.pdf

7. F.B. Chiao, F. Resta-Flarer, J. Lesser, J. Ng, A. Ganz, D. Pino-Luey, H. Bennett, C. Perkins Jr and B. Witek, "Vein visualization: patient characteristic factors and efficacy of a new infrared vein finder technology," Brit. J. Anaesth **110**(6), 966–971 (2013).

8. T.S. Perry, "Veebot: Making a robot that can draw blood faster and more safely than a human can," http://spectrum.ieee.org/robotics/medical-robots/profile-veebot (2013).

9. R.D. Brewer and J.K. Salisbury, "Visual Vein-Finding for Robotic IV Insertion," IEEE International Conference on Robotics and Automation, (2010), pp. 4597–4602.

10. Veebot's Robot Technician Draws Blood From Patients, With Higher Accuracy Rate, http://www.medicaldaily.com/veebots-robot-technician-draws-blood-patients-higher-accuracy-rate-video-248176

11. E. Nakamachi, Y. Morita, and Y. Mizuno, "Development of Automatic 3D Blood Vessel Search and Automatic Blood Sampling System by Using Hybrid Stereo-Autofocus Method," International J. Opt. **2012**, 258626 (2011).

12. A. Chen, K. Nikitczuk, J. Nikitczuk, T. Maguire, M. Yarmush, "Portable robot for autonomous venipuncture using 3D near infrared image guidance," Technology (Singap World Sci) 01 (2013).

13. V.C. Paquit, K.W. Tobin, J.R. Price, F. Mriaudeau, "3D and multispectral imaging for subcutaneous veins detection," Opt. Express **17**(14), 11360–11365 (2009).

14. A. Shahzad, M. Naufal Mohamad Saad, N. Walter, A. Saeed Malik, F. Meriaudeau, "A review on subcutaneous veins localization using imaging techniques," Curr. Med. Imaging Rev **10**(2), 125–133 (2014).

15. M. Zhou, Z. Wu, D. Chen, and Y. Zhou, "An improved vein image segmentation algorithm based on SLIC and Niblack threshold method," Proc. SPIE **9045**, 90450D (2013).

16. L. Zhou, M.S. Rzeszotarski, L.J. Singerman, and J.M. Chokreff, "The detection and quantification of retinopathy using digital angiograms," IEEE Trans. Med. Imaging **13**(14), 619–626, (1994).

17. N. Miura, A. Nagasaka, and T. Miyatake, "Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification," Mach. Vision Appl. **15**(4), 194–203 (2004).

18. M. Vlachos and E. Dermatas, "Vein segmentation in infrared images using compound enhancing and crisp clustering," Volume 5008 of the series Lecture Notes in Computer Science, (2008) pp. 393–402.

19. M. Vlachos and E. Dermatas, "Supervised and unsupervised finger vein segmentation in infrared images using KNN and NNCA clustering algorithms," XII Mediterranean Conference on Medical and Biological Engineering and Computing (2010), pp. 741–744.

20. J. Yang and J. Yang, "Multi-channel gabor filter design for finger-vein image enhancement," Fifth International Conference on Image and Graphics (2009), pp. 87–91.

21. Y. Zhao, L. Rada, K. Chen, S.P. Harding, and Y. Zheng, "Automated Vessel Segmentation Using Infinite Perimeter Active Contour Model with Hybrid Region Information with Application to Retina Images," IEEE Trans. Med. Imaging **34**(9), 1797–1807 (2015).

22. Y. Zhao, Y. Liu, X. Wu, S.P. Harding, and Y. Zheng, "Retinal Vessel Segmentation: An Efficient Graph Cut Approach with Retinex and Local Phase," PLoS ONE **10**(4), e0122332 (2015).

23. W. Kang and F. Deng, "Vein Image Segmentation Based on Distribution Ratio of Directional Fields," Acta Automatica Sinica **35**, 1496–1502 (2009).

24. M.M. Chen, S.X. Guo, and X.H. Qian, "Finger vein image segmentation based on an improved LBF active contour model," J. Jilin University **41**, 1171–1176 (2011).

25. X. Hu, Y. Zhou, and Z. Wu, "A 2.5 dimensional vein imaging system for venipuncture," Proc. SPIE **8668**, 86685A (2013).

26. W.J. Cong, J. Yang, D.N. Ai, Y. Chen, Y. Liu, and Y.T. Wang, "Quantitative Analysis of Deformable Model-Based 3-D Reconstruction of Coronary Artery From Multiple Angiograms," IEEE Trans. Biomed. Eng. **62**(8), 2079–2090 (2015).

27. J. Yang, W.J. Cong, Y. Chen, J.F. Fan, Y. Liu, and Y.T. Wang, "External force back-projective composition and globally deformable optimization for 3-D coronary artery reconstruction," Phys. Med. Biol. **59**(4), 975–1003 (2014).

28. V.C. Paquit, F. Meriaudeau, J.R. Price, and K.W. Tobin, "Simulation of skin reflectance images using 3D tissue modeling and multispectral Monte Carlo light propagation," 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2008), 447–450.

29. Z. Zhang, "A Flexible New Technique for Camera Calibration," IEEE Trans. Pattern Anal. **22**(11), 1330–1334 (2000).

30. Z. Shoujun, Y. Jian, W. Yongtian, and C. Wufan, "Automatic segmentation of coronary angiograms based on fuzzy inferring and probabilistic tracking," Biomed. Eng. Online **9**, 40 (2010).

31. P. Bankhead, C. N. Scholfield, J. G. McGeown, and T. M. Curtis, "Fast retinal vessel detection and measurement using wavelets and edge location refinement," PloS One **7**, e32435 (2012).

32. O. Friman, M. Hindennach, C. Khnel, and H. O. Peitgen, "Multiple hypothesis template tracking of small 3D vessel structures," Med. Image Anal. **14**(2), 160–171 (2010).

33. H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," 9th European Conference on Com-

puter Vision (2006), pp. 404–417.

34. M. Al-Rawi, M. Qutaishat, and M. Arrar, "An improved matched filter for blood vessel detection of digital retinal images," Comput. Biol. Med 37, 262–267 (2007).

1. Introduction

Intravenous injection is a basic and rapid medical method in bedside care. Intravenous therapy can quickly deliver electrolytes, nutrients, and water to the body by inserting a catheter into the veins. Among the 500 million intravenous insertions performed annually worldwide, nearly 14 million failure cases occur on the first attempt [1]. Failed insertions generally cause pain and bruising, and may even lead to venous sclerosis. The primary reason for failure insertions is the difficulty of locating veins precisely based only by virtue of visual and tactile experience of medical professionals. Small veins in women and children, veins beneath dark skin color, or veins in fat layer in obese patients are extremely difficult to visualize or feel by practitioners. Moreover, narrow veins caused by increasing blood pressure resulting from anxiety aggravate the difficulty in identifying veins.

Enhancing the localization of subsurface veins may improve the success rate of intravenous injection [2]. Ultrasound can visualize the blood vessel but with extra assistance and expense. White or red light source may also be transmitted through thin body parts. However, skin burns may be caused by excessive heat production. Non-invasive near-infrared light (NIR) may also be used. High absorption of blood with more deoxygenated hemoglobin for NIR develops the veins much darker than surrounding tissues in the wavelength of 700-900 nm [3]. NIR imaging systems are commercially available to increase the success of phlebotomists, nurses, and anesthesiologists.

VascuLuminator (de Konigh Medical Systems, Arnhem, the Netherlands), developed in [2], was based on transillumination of the puncture site with NIR. The maximum depth of visibility is 5.5 mm for a 3.6 mm blood vessel of phantom data. However, the high intensity of transillumination light needed may cause a burning sensation in opisthenar. Two other devices currently on the market, namely, VeinViewer Vision [4] (Christy Medical Corporation, Memphis, TN, USA) and the AccuVein (AccuVein AV300; Avant Medical, Cold Spring Harbor, NY, USA), use the reflection of NIR light instead of transillumination. VeinViewer Vision allows maximum 10 mm depth of clinically relevant veins to be observed [5].AccuVein provides a permanently alignment, and the center line accuracy of which is less than the width of a human hair [6]. Both instruments are designed as vein contrast enhancers to reveal vessels invisible to the eyes by locating subcutaneous veins and back projecting them on the skin surface. However, the relative depth or diameter of vessels is not provided by both devices. Additionally, VueTek Veinsite [7] (VueTek Scientific LLC, Grey, Maine, USA), a head-mounted vascular imaging system, is commercially designed to offer hands-free and portable operation. Concentrating on automatic intravenous insertion, robotics is combined with image analysis software. The performance of the system is evaluated by counting the number of the emerging veins for patients with different ages, genders, heights, weights, ethnicities and skin colors. Veebot (Veebot LLC) [8] and 7 DOF robotic IV insertion system [9] are developed to achieve automatic vein recognition and intravenous catheterization autonomously. Veebot achieves 83% accuracy rate in selecting the best vein for insertion [10]. 76% of individuals are successfully inserted according to the subjective judgment of medical doctors by 7 DOF robotic IV insertion system. The aforementioned systems do not provide the 3D topography of the surface and vein structures that are significant factors for precise injection. In addition, Doshisha University designed a 3D blood vessel search system, which is installed in a portable self-monitoring blood glucose device [11]. Blood vessel locations are investigated using this device. Only phantom data are evaluated, and bifurcation of vessels is not considered. The results on real vein imaging present

serious errors because of the exclusion of complex blood vessel structures. The portable robot for autonomous venipuncture developed in [12] presents a stereo camera system to construct the 3D vein by moving cameras. The system requires a precise motorized system to control the motion of cameras, thus, the implementation complexity of the system is high. On the other hand, the reconstruction of 3D skin is not described in detail and the depth below the skin surface is not obviously available. The mean completion time is 18.46s for imaging, image processing, and stereo. A 3D imaging system [13] is developed to provide 3D information of the arm surface and automatically detect veins. The system requires a laser stripe line generator for 3D skin reconstruction and no 3D structure of veins is given. The detailed overviews for subcutaneous veins localization using imaging techniques are referred to a recent review [14].

The improvement of subcutaneous vein recognition has received much attention. NIR imaging for subcutaneous veins has generally a lower image quality, and blood vessel contours are relatively obscure. Traditional methods cannot be directly used to segment subcutaneous veins. To date, state-of-the-art segmentation methods for subcutaneous veins are mainly divided into two categories, namely, image patch-based threshold selection segmentation methods and structure feature-based vein segmentation methods. For the first category, Zhou *et al.* proposed a simple linear iterative clustering technique which divides the image into small polygon patches and completes vein segmentation with the NiBlack threshold for each patch [15]. Chen *et al.* [16] proposed a fast local threshold calculation method based on grayscale morphology to realize the local adaptive threshold segmentation. These methods are fast but susceptible to the environment and illumination changes, because only the gray-level distribution, instead of the structural characteristics of blood vessels, is considered. For the second category, Naoto *et al.* introduced a repeated line tracking method based on the characteristic of the gray ridge distribution of blood vessel contours [17] to track for vein recognition. Marios *et al.* combined Hessian matrix and template matching by considering the blood vessel structure to enhance the vein [18], and applied KNN and NNCA for vein segmentation [19]. Yang used a multi-channel Gabor filter-based vein enhancement method [20]. Veins with different angles and widths were enhanced by the Gabor parameter model and reconstructed to obtain the final vein enhancement image. Zhao *et al.* [21, 22] proposed an infinite perimeter active contour model with hybrid region information. Image orientation field [23] uses the distribution ratio of directional fields and a specific threshold to perform vein segmentation. An active contour-based method [24] improves local binary fitting by fuzzy localness to segment veins with inhomogeneous intensity. However, these two methods obtain abundant blood vessel information with high noise in the vein segmentation results.

The specificity and visualizing depth of subcutaneous vein are considered in the development of 3D vein reconstruction systems. These systems commonly assist in intravenous injection and identification of veins to enhance the success rate of intravenous injection further. Hu et al. [25] used edge extraction and the sum of absolute difference matching to realize 3D vein reconstruction and visualize a relatively complete structure. However, this method attained a rough reconstruction of 3D vessels, which was only used to match the structural similarity and could not be applied to 3D localization of venous vessels [26, 27]. Paquit et al. [28] introduced a method to visualize the 3D skin surface and subcutaneous structures using 3D modeling, ray tracing, and Monte Carlo light propagation modeling. However, reconstructing and visualizing veins with a complete structure and accurate positioning are difficult using the existing methods.

Fast and precise guidance to define the position of subcutaneous veins better still needs to be addressed. An imaging system is expected to perform as follows: 1) find vein boundaries with high sensitivity; 2) reconstruct the 3D topography of the surface and veins with high precision; and 3) visualize processed veins intuitively. The imaging system should also be cost effective. Thus, a prototype imaging and augmented system (IAS), which satisfies the aforementioned
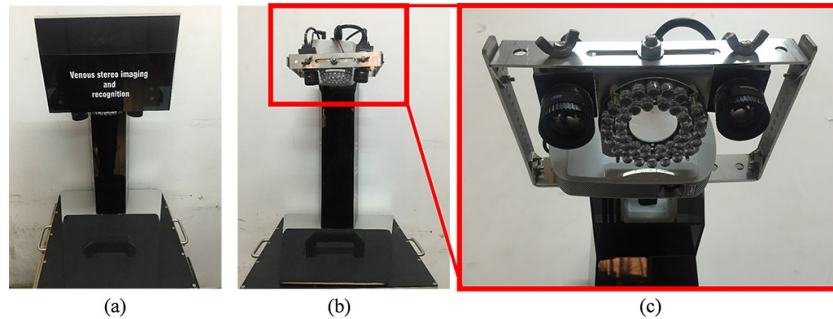
Fig. 1. Prototype of the imaging and augmented system: (a) front view; (b) front view without cover; and (c) enlarged view of the part in the red box.

design requirements, is developed to assist intravenous injection effectively. The advantage of the IAS is based on two inseparable parts, namely, hardware design (imaging system) and software development (image processing). Seven components are integrated to construct the system based on the design scheme, control flow and calibration method. These components include two industrial cameras, a color micro projector, NIR light sources, control circuit, a computer monitor, a support structure with multi-degree of freedom, and an underpan. The following methods are proposed for image processing: 1) a vein segmentation method by combining multi-scale information and structural information; 2) 3D reconstruction of subcutaneous veins according to epipolar constraints and hierarchical optimization, and back-projection on the skin surface for augmented reality; and 3) skin surface reconstruction using active structured light with spatial encoding values, and fusion displayed with the reconstructed vein with indicative vein depth. Thus, we present an improvement of state-of-the-art devices by considering all the requirements for intravenous injection. Figure 1 shows the prototype IAS and will be demonstrated in detail.

The contributions of this work are summarized as follows:

- We build a novel augmented reality system for real-time subcutaneous vein imaging, including system construction and system calibration.

- The special processing methods for the IAS are proposed to segment and reconstruct veins which are further fused with reconstructed skin and projected into the real skin.

- The whole process for subcutaneous vein imaging is realized with a quick and accurate response in real-time.

- The reconstructed veins and the skin are simultaneously visualized, and the indication of vein depth below the skin surface is provided.

2. Imaging system

2.1. System construction

The overall assembly drawing for IAS is shown in Fig. 2. The imaging system is composed of two industrial cameras C_1 and C_2, a color micro projector P, a set of NIR rings of LEDs L, NIR light source control circuit, a computer monitor, a support structure with six-degree of freedom, and a underpan with a hand rail H. In Fig. 2(e), $\Phi_1, \Phi_2, ..., \Phi_6$ denote six-degree of freedom: rotation Φ_1 and Φ_3 of two cameras, translation Φ_2 of the NIR rings, translation Φ_4
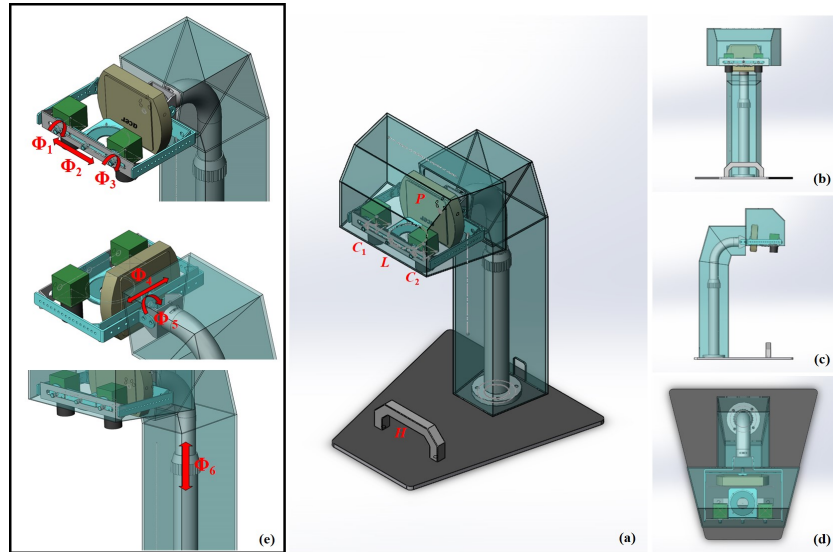
Fig. 2. (a) The overall system assembly drawing; (b) Front view; (c) Left view; (d) over-looking; (e) Six-degree of freedom of the support structure.

and rotation Φ_5 of the projector, and height adjustment Φ_6. The support structure cannot be arbitrarily adjusted when the visual field of patients hand is fixed. L, C_1 and C_2 form the visual system of the 3D reconstruction of subcutaneous vein. P, C_1 and C_2 form the visual system of 3D reconstruction of the skin surface. The computer monitor controls the cameras and projector, system calibration, segmentation and reconstruction of subcutaneous veins, reconstruction of skin surface, 3D data visualization, and vein back-projection.

Industrial cameras are used for imaging both subcutaneous veins and skin surface because three kinds of system design are considered as follows. (1) The crosswise size of the system is reduced with camera multiplexing. (2) The angle between cameras is enlarged by adding a number of cameras for close imaging distance. Thus, the difficulty of time control is reduced and the imaging cross region is expanded with camera multiplexing. (3) The relative position between veins and skin is directly obtained with camera multiplexing. Camera multiplexing requires that the photosensitive range of cameras should simultaneously cover NIR light (700-900 nm) and visible light (450-700 nm). Meanwhile, exposure time and luminance gain should be adjusted in a certain range to guarantee the clarity and contrast of images during mode conversion of imaging. Mini USB cameras DH-HV1351UM-ML with 1.3 megapixels (Daheng IMAVI-SION, China) meet all the above requirements and are used as industrial cameras. Monofocal lenses M1214-MP2 with 12mm fixed focal length (Computar, Japan) are selected for close-up applications required in image processing.

Veins appear much darker than the surrounding tissue at 850nm NIR wavelength. Thus, the 30° Infrared 36-LED Illuminator Board Plate (FY-10036A-30) with peak velocity at 850nm is used as a light source. The LEDs are focused on the target area, that is, the dorsum of the hand, and set in the middle of the cameras to compensate illumination during imaging.

Projecting encoding images and back-projecting reconstructed veins are achieved by a micro projector. Acer C205 with rated brightness of 200 lumens and native WVGA (854×480 pixels) resolution is used as the micro projector and satisfies the requirement of small size ($140 \times 108 \times 27mm^3$) and maximum focal length (7.18 mm).

The light source and cameras are placed on the same side of the target area. The reflected

light by LEDs is captured by the cameras on the same side. The distance between the target area and the cameras is set to 35cm. A control circuit is connected to a computer monitor by a serial port. The circuit is designed to control the alternate conversion between NIR and encoding light sources. The circuit uses MSP430 single-chip microcomputer as control chip, N channel field-effect tube SI2320 A2SHB as the electronic switch, and USB-SERIAL CH340 as the serial port.

The handrail H placed on the underpan limits the hand shapes of patients. Making a fist when holding the handrail could improve the contrast of veins in the NIR images because of tauter skin. Moreover, this practice can reduce the influence of subcutaneous tissues for imaging. The dark substrate of NIR diffuse reflectance is tiled under the handrail to guarantee the gray difference between the background and dorsum of the hand.

2.2. System calibration

The IAS is calibrated to determine the internal and external parameters of the two cameras and the projector. Calibration is performed when the positions of the cameras and projector have been fixed.

(1) Camera calibration

Assuming that the monocular camera is pinhole imaging. We denote a 2D point as $\mathbf{p}_0 = [u,v]^T$ and a 3D point as $\mathbf{p}_W = [X_W,Y_W,Z_W]^T$. After adding 1 as the last element of $\mathbf{p}_0$ and $\mathbf{p}_W$ (i.e. $\mathbf{p}_0 = [u,v,1]^T$ and $\mathbf{p}_W = [X_W,Y_W,Z_W,1]^T$), we express the relationship between the 3D point $\mathbf{p}_W$ and its image projection $\mathbf{p}_0$ by the following calibration model [29]:

$$\lambda \mathbf{p}_0 = \mathbf{A} \begin{bmatrix} \mathbf{R} & \mathbf{T} \end{bmatrix} \mathbf{p}_W \tag{1}$$

where λ is an arbitrary scale factor, $\mathbf{A} = \begin{bmatrix} a_x & 0 & u_0 & 0 \\ 0 & a_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ is the internal parameter matrix,

and $\begin{bmatrix} \mathbf{R} & \mathbf{T} \end{bmatrix}$ is the external parameter matrix. $\mathbf{R}$ and $\mathbf{T}$ indicates the rotation and translation between the world and camera coordinate systems, and includes three parameters $(\theta_1, \theta_2, \theta_3)$ and (t_1, t_2, t_3), respectively. To correct the image distortion caused by the curved lens, the distortion parameters are usually introduced into the calibration model, including the radial distortion parameters $[k_1, k_2, k_3, k_4, k_5, k_6]$ and the tangential distortion parameters $[p_1, p_2]$. The conversion relationship between the imaging and distortion coordinate systems is as follows:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \left(\frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} \right) \begin{bmatrix} x_d \\ y_d \end{bmatrix} + \begin{bmatrix} 2p_1 x_d y_d + p_2(r^2 + 2x_d^2) \\ p_1(r^2 + 2y_d^2) + 2p_2 x_d y_d \end{bmatrix} \tag{2}$$

where, (x,y) is the imaging coordinate, (x_d, y_d) is the distortion coordinate, and $r^2 = x_d^2 + y_d^2$. Moreover, the conversion relationship between the imaging and image systems is

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_x & 0 & u_0 \\ 0 & a_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3}$$

Based on the conversion relationship between different coordinate systems, planar checkboard patterns are used to calibrate the camera by calculating 18 parameters of the camera [29]. These parameters include the internal (i.e., $[a_x, a_y]$, $[u_0, v_0]$, and $[k_1, k_2, k_3, k_4, k_5, k_6, p_1, p_2]$) and the external (i.e., $[\theta_1, \theta_2, \theta_3]$ and $[t_1, t_2, t_3]$) parameters. The external parameters (the number is 6) are the same for each planar checkboard pattern if the world coordinates are fixed. The internal parameters (the number is 12) are the same for all planar checkboard patterns. According
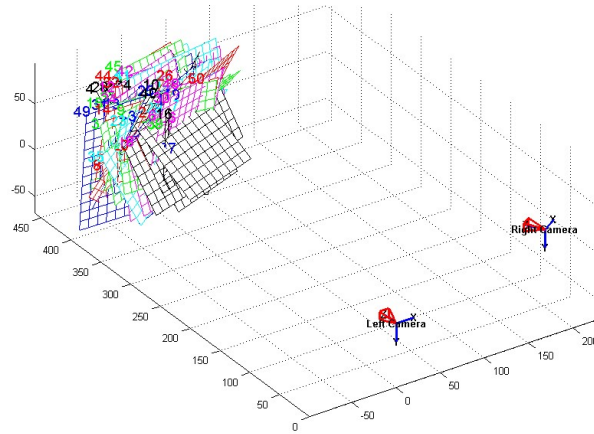
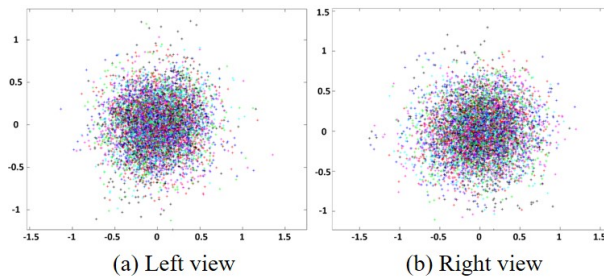Fig. 3. Pose relationship of binocular vision system.

(a) Left view   (b) Right view

Fig. 4. Back-projection error of left and right views.

to Eq. (2), one square corner (x_d, y_d) in the planar checkboard pattern provides two equations. Assuming that K planar checkboard pattern images in different orientations are used to calibrate the camera, and N square corners are extracted for each image. Thus, to calculate all the parameters, the number of $2NK$ should be larger than the number of the parameters of all the images (i.e., $2NK > 6K + 12$). Moreover, only four square corners in a planar checkboard pattern are linearly independent, which indicates that $N = 4$ and $2 \times 4K > 6K + 12 \Leftrightarrow K > 6$. Thus, at least six images are required to ensure the parametric solution. In this study, to reduce the calibration error, 50 planar checkboard patterns in different orientations are collected for an accurate calibration. The simulation diagram for 50 planar checkboard patterns is shown in Fig. 3. Back-projection errors calculated from the left and right views of cameras are shown in Fig. 4.

(2) Projector calibration

The projector without an imaging function can only obtain the image coordinate system without the world coordinate system. The left camera C_1 as a calibrated camera is utilized to access the world coordinate system. The calibration plate for the projector consists of marked points and projection area. The calibration plate is designed as Fig. 5.

The four-step projector calibration is performed as follows. First, relative positions among the projector, calibrated camera, and calibration plate are fixed. Second, marked points (m_1, m_2, m_3, m_4) are then extracted from the calibration plate by the camera, and the external parameters between the camera and world coordinate systems are calculated. Third, the planar
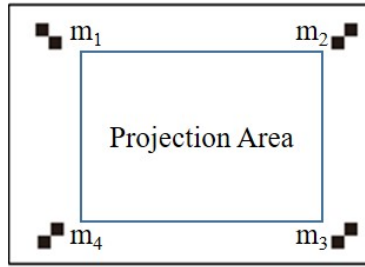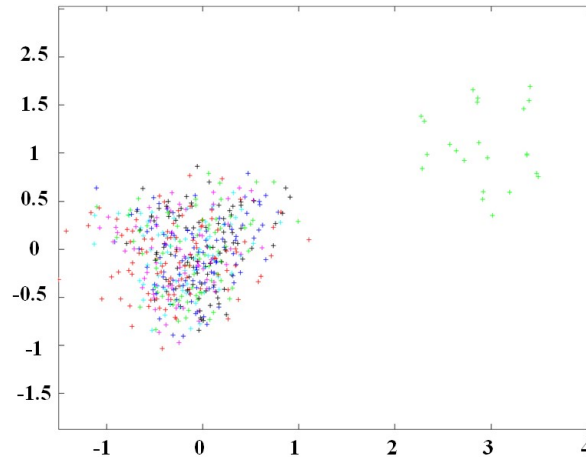
Fig. 5. The projector calibration plate.

Fig. 6. Back-projection error of the projector calibration.

checkboard pattern is projected into the projection area by the projector. The image composed of projected planar checkboard pattern and marked points is taken by the camera. The world coordinate system of the marked points is calculated by external parameters. Finally, linear equations of the internal and external parameters of the projector are set up and solved based on the world and image coordinates of the marked points. According to the internal and external parameters of the projector, the projected planar checkboard patterns are back-projected into image coordinates, and the error of the projector calibration is shown in Fig. 6.

3. Imaging processing

The NIR images have low contrast; thus, the subcutaneous vein images are further analyzed and processed using Microsoft Visual Studio C++ with OpenCV library and Qt for graphical user interfaces after the images are captured. Vein segmentation, 3D reconstruction of subcutaneous veins, and skin surface are implemented to obtain the 3D topological structure of veins and the skin. The processed subcutaneous veins are further back-projected onto the dorsum of patients - hands for augmented veins.

3.1. Vein segmentation

Isotropic undecimated wavelet transform (IUWT) is composed of wavelet decomposition and image reconstruction. Wavelet decomposition gives wavelet scale images $C_i =$
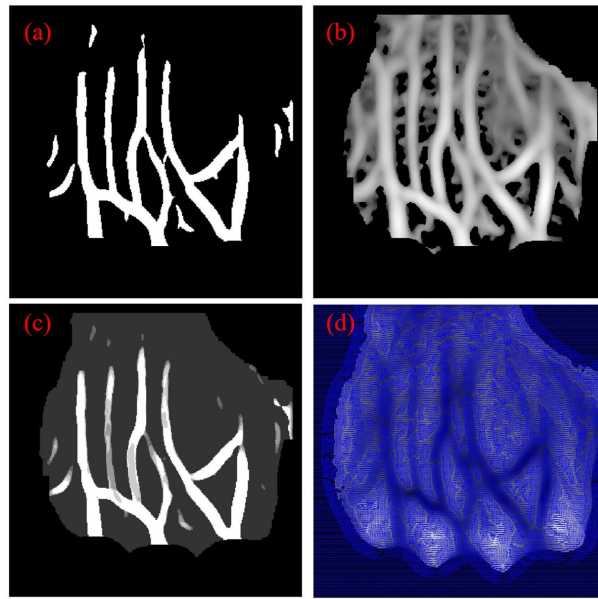
Fig. 7. Feature images: (a) rough segmentation image, (b) vein similarity image, (c) vein scale image, and (d) vein direction image.

$$\begin{cases} I_0 & i=0 \\ C_{i-1} \otimes h_{i-1} & i>0 \end{cases}$$ and wavelet decomposition images $w_{i+1} = C_i - C_{i+1}$. Here, I_0 is the original image, and h_{i-1} is the multiple scale kernel function. The procedure of wavelet decomposition obtains the detail information of the image. The procedure of image reconstruction is a combination of different wavelet decomposition images. Since vascular characteristics are mainly contained within a continuous scale range $[\sigma_1, \sigma_2]$, the reconstructed image $I = \sum_{i=\sigma_1}^{\sigma_2} w_i$ can enhance the multiscale veins: exclusion of small-scale suppress image noise, while exclusion of large-scale correct non-uniform illumination.

Tubular structure analysis is achieved by eigen-decomposition of a Hessian matrix according to the second-order partial derivative of the image. Eigenvalues λ_1 and λ_2, as well as the corresponding eigenvectors v_1 and v_2, are obtained by singular value decomposition of the Hessian matrix and $|\lambda_1| \leq |\lambda_2|$. Four feature images are further obtained as follows:

(a) Rough segmentation image: according to the eigenvalue λ_2, the rough vein segmentation result are obtained by $\max\limits_{\sigma_{\min} \leq \sigma_k \leq \sigma_{\max}} \left(\dfrac{\lambda_2^k}{\sigma_k} \right) > \sqrt{t_2 \pi}$. The threshold t_2 is determined in each scale σ_k.

(b) Vein similarity image: $V = \begin{cases} \max\limits_{\sigma_{\min} \leq \sigma \leq \sigma_{\max}} \ln\left(\lambda_2(I(u,v), \sigma_k)^2 + 1 \right) & \lambda_2 < -\sqrt{t \pi} \sigma_k \\ 0 & else \end{cases}$,

where, $\sigma_{\max}$ and $\sigma_{\min}$ are the maximum and minimum scales of the Gaussian kernel function; t is the filtration factor, which filters the background without destroying the integrity of veins [30].

(c) Vein scale image: maximum scale of each pixel for λ_2^k / σ_k; and

(d) Vein direction image: direction of each pixel denoted by the eigenvector v_1.

According to four feature images shown in Fig. 7, refined segmentation is implemented.

The vein centerlines are obtained by thinning the rough segmented veins in Fig. 7 (a). The endpoints and bifurcation points as the feature points of veins are extracted according to the neighborhood information of points in the vein centerlines. The veins are further split into branches by tracking and going through all the feature points. The vein direction image shown in Fig. 7(d) is further used to repair the structural deficiency in the bifurcation, end of veins, and weak contrast regions. Piecewise spline fitting proposed in [31] is used to simultaneous fit and interpolate the veins to obtain the uniform branch center lines and accurate direction information of the branch. Based on the method proposed in [32], a contour image is generated. 2D features for each pixel are obtained by projecting the vein similarity image Fig. 7(b) and IUWT image Fig. 7(c) into the contour image, which are further clustered into three categories, namely, veins, fuzzy region, and background. The relative categories are projected into the original image space. The refined veins are then finally segmented by filled holes in the blood vessel with morphological closing operation, and the small area in the image is removed with morphological connected-domain method.

3.2. 3D reconstruction of veins

Without loss of generality, veins are assumed to locate in a plane because of the long imaging distance between veins and cameras. Thus, homography, as a weak constraint of the relationship between left/right views, is used to assist vein matching. The aforementioned assumption ensures a fast, real-time and relatively accurate 3D reconstruction of veins. Speeded-up robust features (SURF) [33] are extracted as candidate points from the vein feature image obtained by proportionally fusing the vein similarity image and vein enhancement image. Homography matched relationship between candidate points in the left and right views is fixed with the RANdom SAmple Consensus (RANSAC) method, and homography-matched points are obtained. The merit of RANSAC is that the model parameters are robustly estimated and not affected by outliers.

The strict constraint of relationship between the homogeneous coordinates of the left and right views is described by the polar line constraint equation. Assuming that points p_1 and p_2 are in the vein centerlines of the left and right views, respectively, we limit point p_2 in the corresponding pole line of p_1 to find the matching point p_2 for p_1. For points in the left view, the candidate matching points in the right view are obtained by the polar line constraint. This result indicates that the intersection between the polar line of p_1 and centerline in the right view denotes the candidate matching points. The matching points are screened with matching difference degree D between homography-transformed points and candidate matching points. The matching difference degree D is defined by combining Euclidean and Cosine distances, as follows:

$$D\left(p^H, p'\right) = \frac{Dis\left(p^H, p'\right)}{\left|\cos\left(\vec{v}^H, \vec{v}'\right)\right|} = \frac{\sqrt{\left(x^H - x'\right)^2 + \left(y^H - y'\right)^2}}{\frac{\left|\vec{v}^H \vec{v}'\right|}{\left\|\vec{v}^H\right\|\left\|\vec{v}'\right\|}} \tag{4}$$

where p^H and $\vec{v}^H$ denote the coordinate and vein direction of homography-transformed point of p_1, respectively. p' and $\vec{v}'$ denote the coordinate and vein direction of the candidate matching point of p_1, respectively. If the coordinate (p^H and p') and vein direction ($\vec{v}^H$ and $\vec{v}'$) are similar, the matching difference degree D is small. The strict matching of the vein centerline is achieved by estimating the matching difference degree D and used to optimize further the homography matrix calculated with SURF matching points. The candidate matching points in the complex

regions are screened based on the known vascular segments to correct the vein centerlines matching. The matching difference degree D between vascular segments is calculated with known centerline matched points and then used to screen the un-matched points. For matched points q_1 and q_2, the relationship between the image coordinate and corresponding 3D space coordinate X is based on a binocular vision imaging model and shown as follows:

$$\begin{aligned} q_{1,i} &= P_1 X_i = K_1 \left[I | 0 \right] X_i \\ q_{2,i} &= P_2 X_i = K_2 \left[R | t \right] X_i \end{aligned} \tag{5}$$

$$AX = \begin{bmatrix} u_1 p_1^{3T} - p_1^{1T} \\ v_1 p_1^{3T} - p_1^{2T} \\ u_2 p_2^{3T} - p_2^{1T} \\ v_2 p_2^{3T} - p_2^{2T} \end{bmatrix} X = 0 \tag{6}$$

where P_1 and K_1 are the projection and internal reference matrices of camera C_1, whereas corresponding matrices for camera C_2 are denoted by P_2 and K_2. $[R|t]$ is the rotation and translation matrix, and p_i^{jT} denotes the jth row of the projection matrix P_i ($i = 1,2; j = 1,2,3$). The 3D coordinates of vein centerlines are attained using the relationship between matched points.

The 2D vessel fitting method [33] is extended in our system to ensure the continuity and smoothness of vein centerlines. The vein centerline is divided into equal segments. The polynomial fitting is implemented between different segments, while, three-spline interpolation is executed in each segment. The number of the segments is determined by the trade-off between the bendability and smoothness of veins. In the vein segmentation procedure, the 2D diameter information is obtained by left and right views, respectively. This information is used to estimate the vein radius and assist the 3D vessel wall reconstruction. The vein radius are processed by polynomial fitting and interpolation. Vein surfaces are obtained by generating cylindrical ring along the tangent direction of 3D vein centerlines, which are shown in Fig. 8.
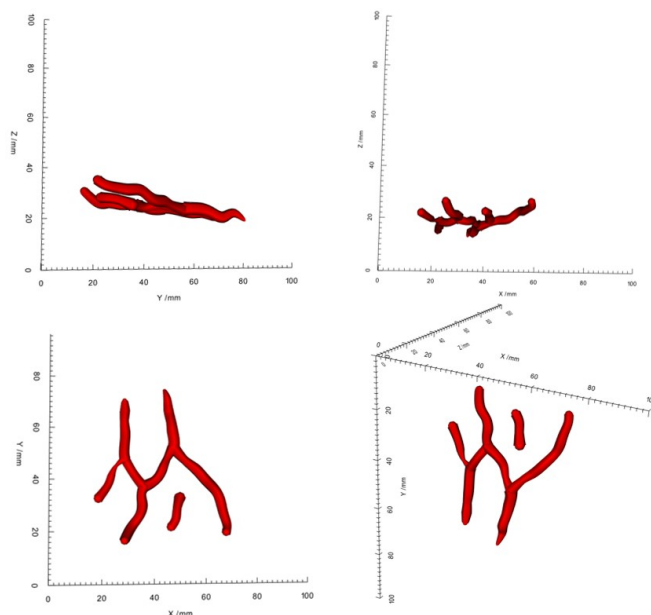

Fig. 8. Surface rendering of 3D reconstructed vein structures with scales in four different views using a real image.

3.3. 3D reconstruction of skins

Skin surface is difficult to reconstruct densely under passive imaging or uniform illumination imaging without a large number of feature points. Thus, the space binary encoder based on Gray code is used for skin surface encoding and generates adequate feature points. Assuming $FuncG(x)$ transforms binary code x into the corresponding Gray code. For an image of size $D_1 \times D_2$, the row/column encoding is as:

$$Gray(j) = FuncG\left(\frac{ps+j}{\mu}\right) \tag{7}$$

where, phase shift $ps = \lfloor (2^n - D_i)/2 \rfloor$, $n = \lceil \log_2(D_i) \rceil$, $i \in \{1,2\}$, D_i is dimension of the image; $Gray(j)$ is the jth row or column. The system projects structure light encoding patterns on the back of the skin surface according to the fixed time sequence, which is shown in Fig. 9. The deformation of the projected lines is due to the spatial variation of opisthenar.

Decoding is the process of determining the column/row encoding values on the area of opisthenar, including judgment of opisthenar area, pixel decoding, and code-word combination. The opisthenar area is judged by calculating the difference between the positive (I_G^+) and negative (I_G^-) code images. The decoding in the opisthenar area is limited by placing a dark diffuse substrate under the handle and the adjusting exposure time of imaging device. This step ensures that evident black and white flips arise only on the hand area. Pixel decoding is performed to solve the encoding value of each pixel using an image sequence. The ith Gray encoding value of a pixel p in the image sequence $G_i(p)$ is calculated by a set of (I_G^+) and (I_G^-). If the difference

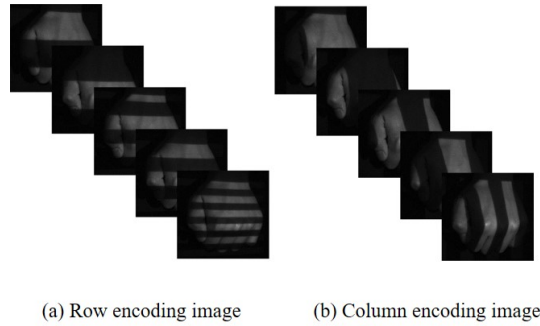(a) Row encoding image   (b) Column encoding image

Fig. 9. Row/column Gray encoding project on the back of the skin surface.
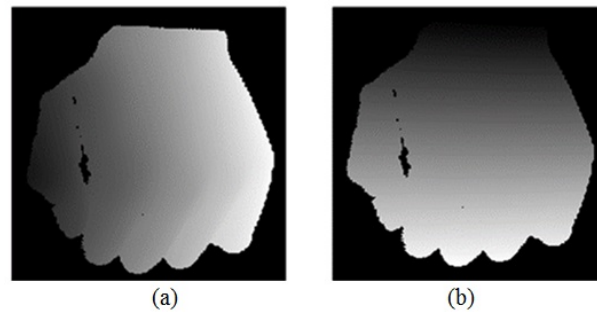

(a)   (b)

Fig. 10. Decoding images obtained by transforming the Gray code into binary values: (a) Column decoding image in the left view; (b) Row decoding image in the left view.

Fig. 11. Surface rendering of 3D reconstruction skin in three different direction using a real hand.

between I_G^+ and I_G^- is more than a default threshold, $G_i(p)$ is equal to one. Otherwise, $G_i(p)$ is equal to zero. Pixels with same code value may be in the adjacent area because of the low contrast of the image. Thus, the center coordinate of the adjacent area is re-calculated as the coordinate of the code value. The decoding image is obtained by transforming the Gray code into binary values, and the results are shown in Fig. 10.

Dense point cloud reconstruction of skin surface is implemented according to the decoding images captured from the binocular cameras. For a point in the left view, we find the corresponding matched point with the same decoding value in the right view using the Hash structure. The skin surface is reconstructed according to the principle of triangulation with all matched point sets. The 3D reconstruction of the skin surface is illustrated in Fig. 11.

4. System performance

A sequence of experiments is conducted to evaluate the performance of the developed IAS for vein segmentation, 3D reconstruction of veins and skins, and augmented display.

4.1. Real vein segmentation

The real subcutaneous vein image is processed to investigate the robustness of the proposed segmentation method. The contrast and acutance of edges are low in the subcutaneous vein image shown in Fig. 12, especially in areas a, b, and c. The contrast of veins in area a is influenced by the light shadow. The veins in deep-seated vessels in area b are almost invisible compared with the superficial-seated vessels. The vein edges in the area c are extremely obscured because of the intersection of multiple vessels.

The vein segmentation results shown in Fig. 13 are obtained using the proposed method and Gaussian matched filters (GMFs) [34]. The proposed method accesses the veins with complete
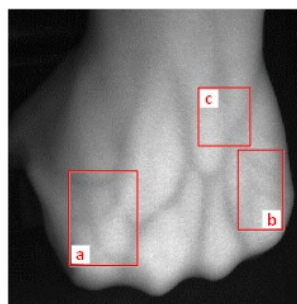

Fig. 12. Real subcutaneous vein image with low contrast and acutance of edges: (a) influenced by the light shadow; (b) deep-seated vessels; (c) obscured intersection of multiple vessels.

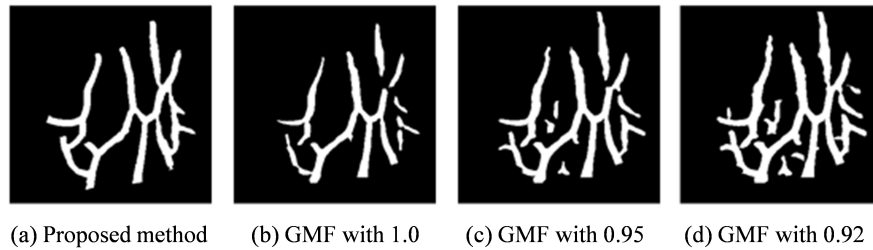(a) Proposed method   (b) GMF with 1.0   (c) GMF with 0.95   (d) GMF with 0.92

Fig. 13. The segmentation results of real subcutaneous vein image by using (a) the proposed method; (b d) the Gaussian matched filtering method of three thresholds 1.0, 0.95 and 0.92.

structure, smooth edge, and precise diameter information. By contrast, GMF cannot adapt for both vein structure and noise interference because the threshold varies in different images.

4.2. Accuracy of vein matching

Accuracy of vein matching is an important measure of successful vein 3D reconstruction. Ground truth of the structure of real subcutaneous veins is unavailable, so the vein back-projection error is applied to describe the accuracy of vein matching. Vein back-projection denotes the vein image that is obtained by projecting the 3D reconstructed vein structure into the imaging plane according to the internal and external parameters of cameras $[\mathbf{R} \ \mathbf{T}]$. The back-projection error is mainly caused by the segmentation accuracy, vein matching accuracy and the internal/external parameters of cameras. Assuming that the proposed segmentation method could obtain the ground truth segmented veins for detecting the accuracy of vein matching. Moreover, a precise system calibration is implemented. Thus, the larger vein matching error reflects the higher back-projection error.

Five different groups of real subcutaneous veins are imaged in the left and right views and shown in the first two rows (a1~a5 and b1~b5) in Fig. 14, respectively. These five sets of vein images with various characteristics are collected from different subjects studying in Beijing Institute of Technology. The ground truth segmented veins is shown in the middle two rows (c1~c5 and d1~d5) in Fig. 14, respectively. According to the proposed 3D vein reconstruction method, the 3D vein structure is obtained and back-projected into two imaging planes, shown in the last two rows (e1~e5 and f1~f5) in Fig. 14, respectively. The red lines in the back-projection images denote the segmented veins without corresponding matching results and cannot be reconstructed. The black areas are the background. The pink, white, and yellow lines denote the veins that are segmented and matched with different back-projection errors. The matching error gradually increases from pink to yellow. The vein structures are evidently sparse and simple in Figs. 14(c1, d1) and (c4, d4). The segmented vein structures in the left view exhibit clear corresponding structures in the right view. Thus, vein matching can be easily achieved using the proposed method. However, the vein structures in left-right views in Fig. 14(c2, d2), (c3, d3) and (c5, d5) are more complex with evident difference. Not all branches of blood vessels in one view can capture the corresponding matching vein in the other view. For these images, the proposed matching method can suppress the disturbance of mismatched blood vessels to complete vein matching. The back-projection error in Figs. 14(e2, f2) and (e3, f3) increases in the area of the end of the opisthenar close to the elbow. This finding is due to the fact that the elbows do not meet the hypothesis of planes, and have large parallax between left-right views. Furthermore, back-projection errors become large near the intersection of blood vessels.

The distribution of back-projection errors is shown in Fig. 15. The mean back-projection errors for the 3D reconstructed veins in the left and right views show the back-projection errors
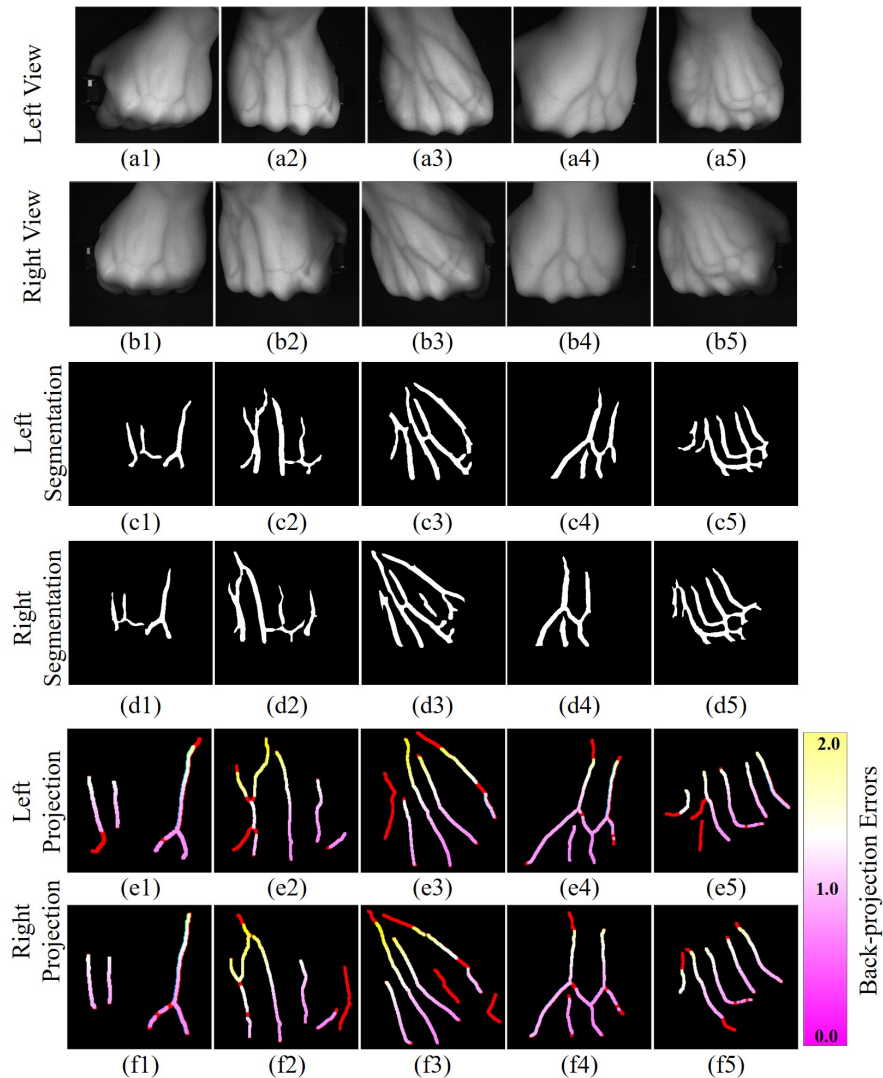
Fig. 14. Vein back-projection error. (a1~a5, b1~b5) Five different groups of real subcutaneous veins are imaged in the left and right views; (c1~c5, d1~d5) The ground truth segmented veins; (e1~e5, f1~f5) Obtained 3D vein structure back-projected into two imaging planes.

produced in the binocular view system. The back-projection errors for 75% of matched points are less than one pixel. The maximum back-projection errors for the Fig. 14(e1, f1), (e4, f4) and (e5, f5) are less than one pixel. Even the end of the opisthenar does not satisfy the hypothesis of planes for Fig. 14(e2, f2) and (e3, f3), the maximum back-projection errors are less than two pixels. Thus, the proposed vein matching method is proved to be accurate.

4.3. Distance error of feature points

2D image processing software CorelDraw is used to automatically measure the distances between different feature points in the simulated vein image, which are indicated as standard
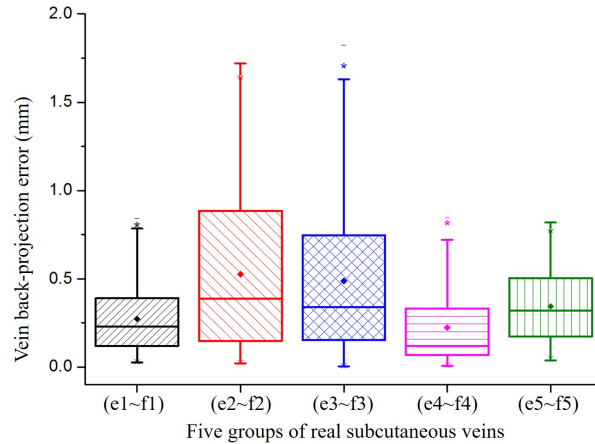
Fig. 15. Back-projection error of vein center lines.

distances. Meanwhile, 3D point cloud processing software MeshLab is used to manually measure the corresponding distances between feature points in 3D point clouds, which are as reconstructed distances. The bifurcation points of veins are selected as feature points. Distances between corresponding feature points in 3D space and 2D image are shown in Fig. 16.

Nine simulated vein images are designed to evaluate the distance error. 10 distance error distribution are measured for each image and shown in Fig. 17. In Fig. 17, all the mean distance errors are less than 1mm, and 75% distance errors are less than 1.5mm.

4.4. Duration times

The completion augmented reality imaging is implemented on single-core single-thread Intel(R) Core(Tm) i7-4720HQ CPU @ 2.6GHz. The mean times of augmented reality back-projection across all 50 trials is 579ms. Figure 18 shows the comparison of duration times with [12] for real-time capability based steps, including vein segmentation, 3D vein reconstruction and vein back-projection. The real-time back-projection overcomes the latency between a movement of the patient hand and the projected vein image updating to reflect that movement. And it effectively assists doctors for locating the injection veins. On the other hand, the projector resolution is of 1280×720. Eleven (i.e. $\lceil \log_2 (1280) \rceil$) Gray coded images are required to
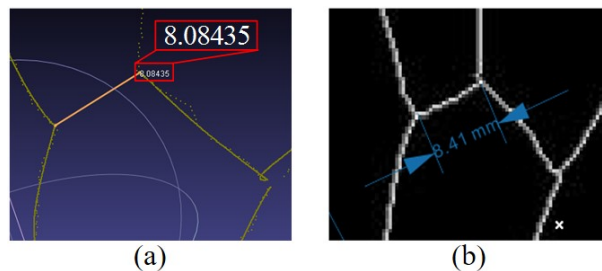

Fig. 16. Measurement for distance between corresponding feature points in 3D space and 2D image. (a) distance between feature points in 3D point clouds; (b) distance between feature points in the simulated vein image.
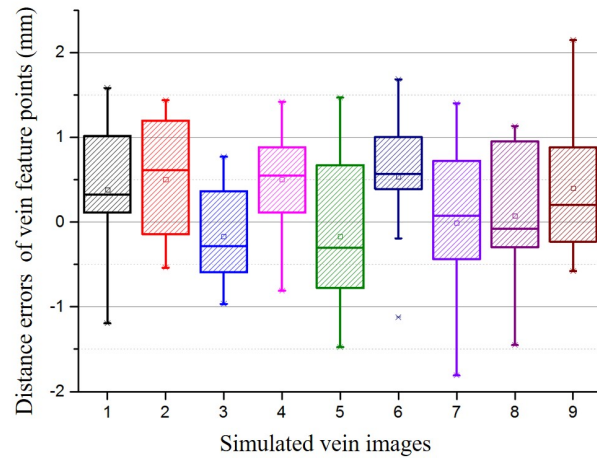
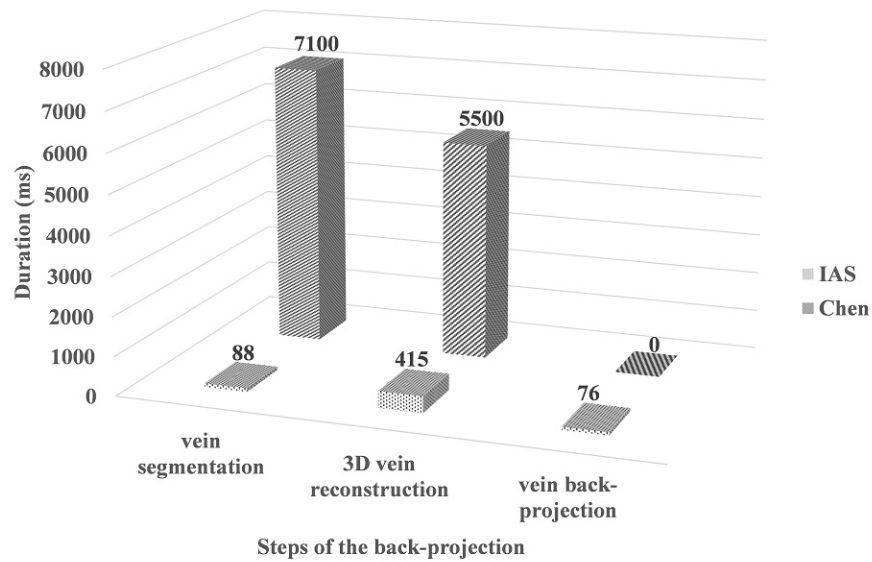Fig. 17. Distance errors of vein feature points for nine simulated vein images.

Fig. 18. Comparison of duration times for each real-time step.

label each projector column and ten (i.e. $\lceil \log_2 (780) \rceil$) for each projector row. This is a total of 21 images which would require 6.7s to acquire on 3D skin reconstruction. The visualization of fusion display takes only 116ms. Thus, 3D fusion display is total 6.81s, which gives intuitive guidance for vein depth below the skin surface and not real-time.
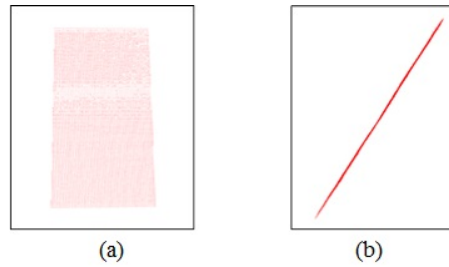
Fig. 19. 3D reconstruction of a plane. (a) flat top view; (b) plane side view.
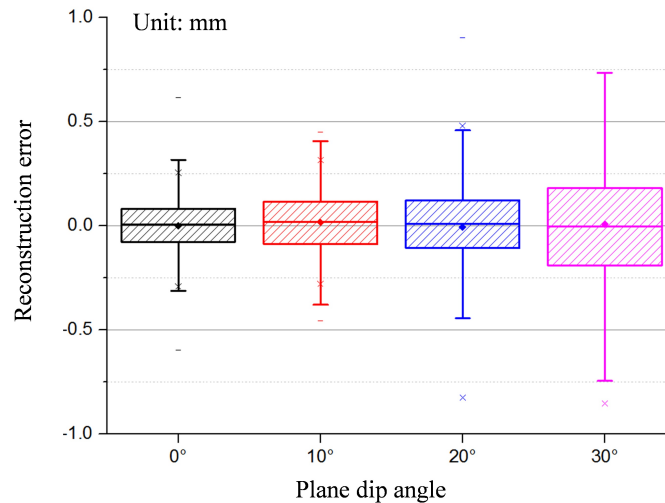

Fig. 20. Relationship between the plane dip angle (0°10°20°30°) and the reconstruction precision of structure light.

4.5. Accuracy of skin reconstruction

The main parameters of the 3D reconstruction of skin surface include structure light encoding parameter (*PixelsPerCode*) and decoding threshold (*Th*). The *PixelsPerCode* determines the encoding spatial resolution. Large *PixelsPerCode* indicates that a single code value corresponds to a large area, spatial encoding is sparse, and skin point cloud is also sparse. By contrast, excessive dense spatial encoding may cause encoding error caused by limited camera resolution. Experimentally, *PixelsPerCode*=2 by considering the projection-range/resolution of the system and the imaging-range/resolution of the cameras. Around 6000 points are rebuilt for the dorsum of the hand. *Th* is varied and related to brightness of the projector, gray scale of projection images, camera exposure time, and camera gain. We set Th=50 to suppress noise interference.

The accuracy of skin reconstruction is evaluated by analyzing the 3D reconstruction of structure light with space plane measurement method. The 3D reconstruction of a plane is shown in Fig. 19, including flat top and plane side views. Gray code-based 3D reconstruction of structure light is evidently effective for reconstructing of space plane.

Meanwhile, the relationship between the precision of structure light reconstruction and the plane dip angle is investigated and shown in Fig. 20. The precision of point cloud reconstruction is less than 0.25mm using Gray code-based 3D reconstruction of structure light. Meanwhile,
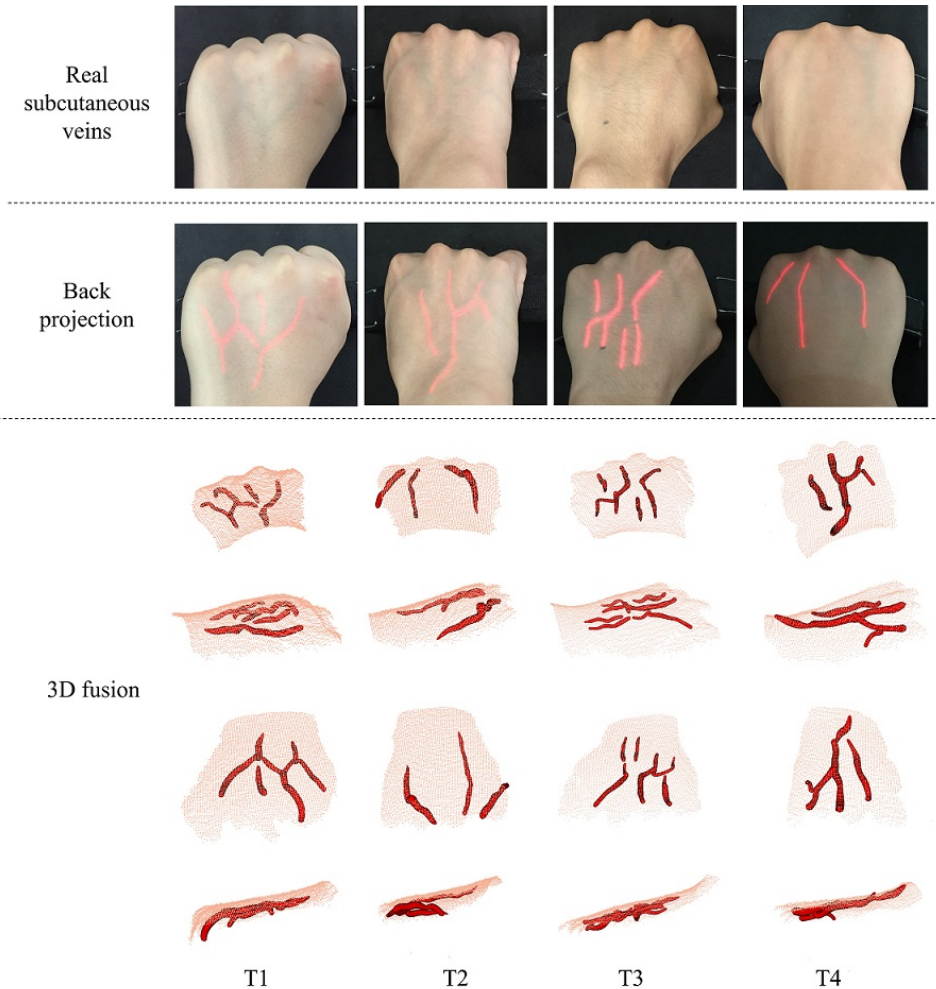
Fig. 21. Back-projection of the real subcutaneous veins. The first row shows four real sub-cutaneous veins; the second row shows the corresponding back-projection images; the last four rows are the 3D rendering and fusion display of the subcutaneous veins and skin surface.

the precision of point cloud decreases with increasing plane dip angle.

4.6. Augmented display

Figure 21 shows four sets of back-projection results of testers by using IAS. The first row in Fig. 21 shows four real subcutaneous veins. The veins are difficult to be identified by eyes. The corresponding back-projection images are shown in the second row in Fig. 21. It is obvious that IAS can find and back project the reconstructed structures into the skin surface to enhance the veins, which effectively assists doctors for locating the injection veins. Three practitioners are employed to evaluate with real veins. According to the opinions of the practitioners, IAS has a superior and fast imaging. Based on the calibration process of the cameras and projector, the 3D reconstructions of veins and skin surface are projected into the same coordinate on the basis of the system calibration parameters. The 3D rendering and fusion display of the subcutaneous

veins and skin surface are shown in the last four rows in Fig. 21.

5. Conclusion and discussion

We have presented a novel 3D reconstruction and fast imaging system for subcutaneous veins by augmented reality. This study is performed to reduce the failure rate and time required for intravenous injection by providing augmented vein structures that are back-projected and superposed on the skin surface of the hand. IAS is not going through the costly and time-consuming process of alignment between the real veins and the back-projected images. Meanwhile, an early evidence of the potential to improve clinical efficacy is provided by giving the fusion display of vein depth below the skin surface. The subcutaneous vein is captured by two industrial cameras with extra reflective NIR lights. A local adaptive subcutaneous vein segmentation method based on multiple feature clustering was proposed to overcome the problems of inaccurate segmentation in bifurcations and tails of veins. The proposed segmentation method provides better peripheral veins, bifurcation structures, and vein edges than those obtained through the Gaussian-matched filter method. Afterwards, the vein structures that are captured in the two cameras are matched and reconstructed based on a novel vein matching and 3D reconstruction method, and back-projected on the skin surface. It overcomes the matching problems caused by the incomplete structure of vein segmentation and asymmetric structures of veins in the left and right views based on the epipolar constraint and homographic property. The skin surface is reconstructed using active structured light with spatial encoding values and fusion displayed with the reconstructed vein. The 3D reconstruction of the skin is designed to limit the reconstruction error within 0.25 mm. The system implements time-sharing multiplex between NIR and visible imaging to reduce the number of required cameras by utilizing the spectral response characteristics of cameras.The system performance has been evaluated with regard to real vein segmentation, vein matching accuracy, distance error of feature points, duration times, skin reconstruction accuracy, and augmented display. All experiments were validated with sets of real vein data. The construction system produced good 3D fusion display results for both subcutaneous veins and skin, even for imaging with low contrast and ambiguous borders. The developed IAS increases the success rate and efficiency of intravenous injection.

In future studies, reflection imaging will be replaced by transmission-reflection imaging using time switching to improve the image quality. Transmission imaging produces the major structure of the vein, while details of the vein structure and the dorsum of the hand are extracted using reflective imaging. Multi-thread will be employed to accelerate the speed of the system and achieves real-time 3D reconstruction of the skin. Furthermore, the intersection of veins will be investigated more intensively to enhance the details of images.