Combining near-infrared illuminants to optimize venous imaging

Vincent Paquit[a,b,c], Jeffery R. Price[a], Fabrice Mériaudeau[b,c], Kenneth W. Tobin[a] and Thomas L. Ferrell[c]

[a] Oak Ridge National Laboratory†, P.O. Box 2008, Oak Ridge, TN 37831, USA;
[b] Le2i, University of Burgundy, 12 rue de la fonderie, 71200, Le Creusot, FRANCE
[c] Department of Physics and Astronomy, University of Tennessee, Knoxville, TN 37996-1200, USA

ABSTRACT

The first and perhaps most important phase of a surgical procedure is the insertion of an intravenous (IV) catheter. Currently, this is performed manually by trained personnel. In some visions of future operating rooms, however, this process is to be replaced by an automated system. We previously presented work for localizing near-surface veins via near-infrared (NIR) imaging in combination with structured light ranging for surface mapping and robotic guidance. In this paper, we describe experiments to determine the best NIR wavelengths to optimize vein contrast for physiological differences such as skin tone and/or the presence of hair on the arm or wrist surface. For illumination, we employ an array of NIR LEDs comprising six different wavelength centers from 740nm to 910nm. We capture imagery of each subject under every possible combination of illuminants and determine the optimal combination of wavelengths for a given subject to maximize vein contrast using linear discriminant analysis.

Keywords: NIR venous imaging, medical robotics, segmentation and rendering, image-guided therapy

1. DESCRIPTION OF PURPOSE

The ultimate goal of our work is to develop an image-guided robotic system for automated catheterization. In previous work,[1] we described benchtop hardware and algorithms to image subcutaneous veins and map the surface topography of the arm or wrist. This system employed NIR illumination for venous imaging and an NIR structured light ranging system for surface mapping. Our purpose in the work described here is to determine the optimal NIR wavelengths for optimizing vein contrast and to determine if and how those optimal wavelengths vary between subjects with different physiological characteristics (e.g., skin tone and/or hair).

2. METHODS

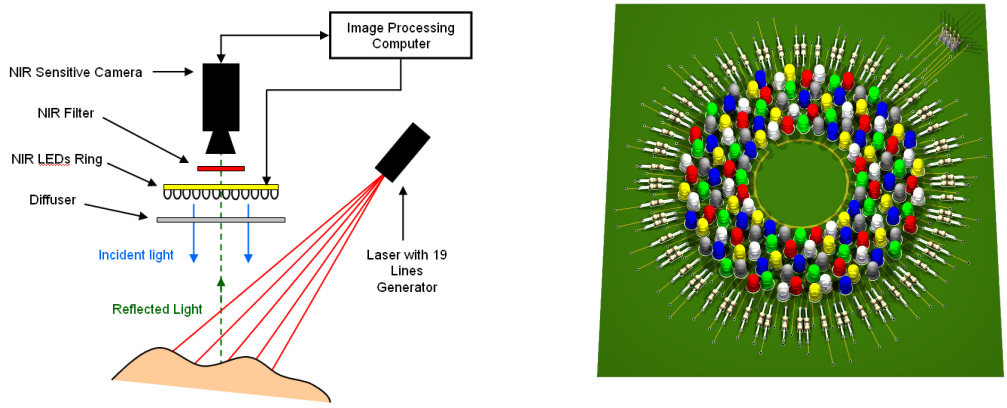
Previous research[2] has investigated the propagation of light in tissue as a function of wavelength and, to some extent, the variation in venous image contrast as a function of (monochromatic) illumination wavelength in the NIR range.[3] We have constructed a benchtop optical system, as illustrated in Fig. 1(a), that allows us to illuminate an area of the forearm or the hand with any combination of the six different wavelength-centered NIR LEDs. For each subject, we acquire 63 NIR-illuminated images, at video frame rates, corresponding to all possible combinations of the six wavelengths ($2^6 - 1$). We also acquire an RGB image of the skin surface with a color calibration target to quantify skin tone and visualize any surface hair. For this experiment, apparent veins are identified either manually or automatically from the NIR and/or RGB images. Each vein and background pixel is then represented as a 63-dimensional feature vector corresponding to the (normalized) intensity under each of the 63 illumination conditions. To find the optimal linear combination of illuminants, we perform two-class linear discriminant analysis (LDA)[4] for the vein vs. background problem. Vein segmentation and identification is based on a line detection algorithm developed in remote sensing for road extraction.[5]

Further author information: (Send correspondence to V.P.)
V.P. : E-mail: paquitvc@ornl.gov, Telephone: 1 865 241 3302

†This paper was prepared by the OAK RIDGE NATIONAL LABORATORY, Oak Ridge, Tennessee, 37831-6285, operated by UT-BATTELLE, LLC for the U.S. DEPARTMENT OF ENERGY under contract DE-AC05-00OR22725.

2.1. Experimental setup

Our acquisition system (Figure 1) is composed of a NIR video camera with a 740nm long-pass interferometric filter, a NIR multi-wavelength ring of LEDs and a NIR line-generating laser module.



(a) Schematic of the acquisition system

(b) 3D rendering of the ring of LEDs.

Figure 1. Schematic of the acquisition system. On Figure 1(b), each of the six wavelengths is displayed with a different color to illustrate the placement.

We have selected a Mutech Phoenix PC-1280/M Monochrome Camera with 10 bit pixel format with a resolution set to 640×480 (VGA) at 60 fps. Our choice was based on the appropriate spectral response of the CMOS sensor in the NIR range of light, its compactness and the communication protocol allowing both video signal transfer and hardware configuration. The camera is connected and controlled by an IBM compatible personal computer (PC) via a USB 2.0 connection.

The optical system is composed of a 25mm lens with a NIR interferometric filter attached. In our setup the NIR filter is used as a high pass filter; combined with the spectral response of the camera we obtain a band pass filter which allows a sensitivity in the spectral range going from 740nm to 960nm.

The lighting source is composed of 144 LEDs mounted on a ring with 24 LEDs for each of the six wavelengths (740, 770, 810, 840, 880 and 910). Each LED has a power dissipation around 30mW. The position of the LEDs on the ring has been determined to get a uniform illumination of the scene even by switching or combining the different wavelengths. In addition we have placed a diffuser in front of the LEDs to increase this uniformity. We use a Measurement Computing PMD-1024HLS 24-line digital I/O module to select a wavelength or a combination of wavelengths for illumination.

We use a 19-line generating beam splitter mounted on a 785nm laser diode to generate the structured light source needed to create the 3D triangulation system. We can cover a surface up to eight square inches with the 19-line generator, therefore providing flexibility and control of the robotic catheter insertion system.

2.2. Data Collection

In this study, we have collected image data of subjects' arms and/or wrists to build a large database of images. Our goal is to find a possible generalization procedure in finding subcutaneous vessels using NIR imaging, which implies to have the most representative panel of the population regardless ethnicity, age, and sex. Following the experimental protocol which was approved by the Oak Ridge Site-wide Institutional Review Board, each subject places a skin surface area from the region going from the extremity of the hand to the antecubital fossa in the system cameras field of view and holds arm still for approximately ten seconds. During this period, the target area will be illuminated by laser-line module (structured light source) and a series of combinations of near-infrared LEDs. The combinations comprise single wavelengths and combinations of single wavelengths. For each unique combination in the series, an image is acquired by the system camera and recorded in the database.

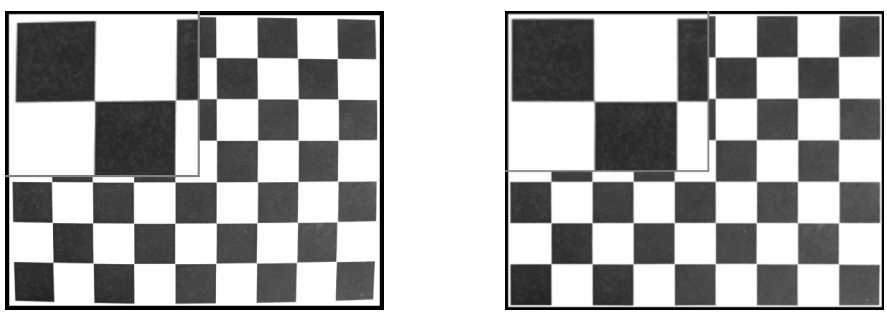
This process has been repeated for different arm/wrist regions.
For each individual we are recording the following information in the database:

1. age, gender, and ethnicity;
2. physical particularities: hairs, birth marks, scars, etc;
3. a digital color picture of the subjects arm/wrist using a standard color digital camera. This picture is taken in presence of a color calibration target, in our case a MacBeth chart, to quantify skin tone for future comparisons;
4. a NIR image of the laser line projection on the surface of the skin. This picture will be used for the 3D reconstruction process and to create a region of interest mask.
5. 63 NIR images of the target area corresponding to the $2^6 - 1$ combinations of the 6 different wavelengths.

2.3. Acquisition System Calibration

2.3.1. Distortion Correction

In computer vision, the calibration of an optical system is used to determine the intrinsic and extrinsic parameters of a camera.[6] These parameters correspond to the mathematical relations between the 3D space of reference and the 2D perception of the scene, with the resulting equations, we are able to correct the distortion of the picture and to compute the metric value of a pixel. Different methods of calibration already exist and many tend to use the same process[6, 7]; we used the method describe in [8] for our application. The calibration is made when the position of the laser compared to the camera and the focal plane of the optical system have been fixed. The method consists of taking pictures of different orientations of a planar checkerboard pattern considered as a metric reference for our system. Notice that we have to illuminate the scene with the NIR lighting source due to the presence of the interferometric filter. The analysis of the position of all the square corners in a set of six to ten pictures give us the intrinsic parameters of the camera: focal length, principal point, skew coefficient, radial and tangential distortions. The Figure 2 presents an example of the correction result obtained with these parameters.



(a) Distorted image of the pattern (b) Undistorted image

Figure 2. Distortion correction example

2.3.2. 3D Modeling

The 3D reconstruction process uses active optical triangulation by combining a camera and a laser line generator.[9] The geometry of the laser triangulation is shown in Figure (3). The camera is aligned along the Z axis and the laser line generator is positioned at a distance b from the camera with the angle θ relative to the X axis. Assuming that the considered laser point coordinates (x, y, z) in the 3D baseline has a projection (u, v) on the image plane, the similar triangles equations give the mathematical relation between the measured quantities (u, v, θ) and the coordinates (x, y, z):

$$[x, y, z] = \frac{b}{f.\cot\theta - u}[u, v, f]. \tag{1}$$

The calculation of the equation of each of the lines during the calibration process depend on the previous triangulation equation. Because the position of the laser compared to the camera is fixed, the 19 equations are calculated only one time. The second part of the 3D reconstruction process depends on the detection of the center of the laser stripe lines. The reflection of the laser light on the surface of the skin is comparable to a blurry Gaussian signal. To detect with accuracy the position of the center of the laser line, we have previously tested with success different subpixel peak detection algorithms such as: Blais and Rioux,[10] Forest[11] and the center of mass. However, the algorithm initially implemented for the detection of the veins and explained in section 3.3 appears to give better detection results and simplify the classification of each of the center line pixels. This classification is performed to associate those pixels with one of the nineteen laser lines equations previously computed and used for the 3D reconstruction (Figure 4).

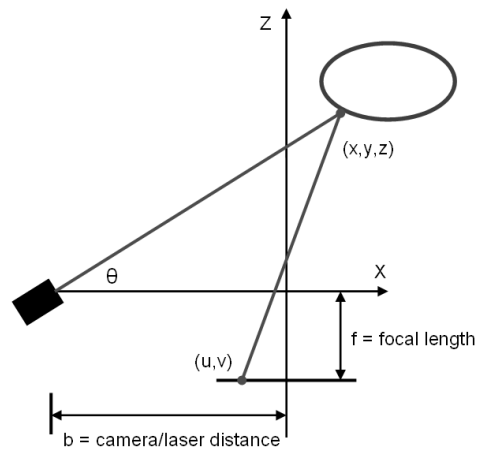


Figure 3. Principle of 3D reconstruction with active optical triangulation

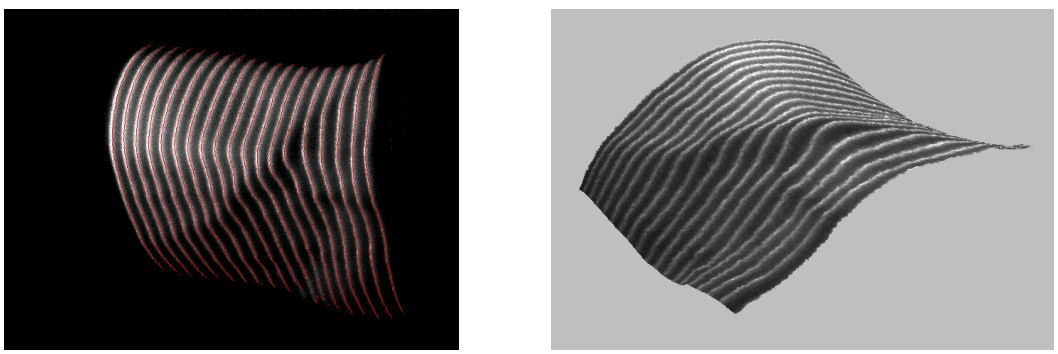


Figure 4. 3D reconstruction visualizations: from 2D to 3D

2.3.3. Spectral Reflectance

By nature each element receiving energy from a complex light source will absorb a portion of the light source and will reflect or transmit the other part as an electromagnetic energy. The spectral response of this element corresponds to the measure of the energy reflected at specific wavelengths. This response characterizes the composition of the element, its state and represents its spectral signature. The physical quantity representing the spectral response is the spectral reflectance which is the reflectance measured within a specific wavelength interval. The reflectance corresponds to the ratio of the radiant energy reflected by a body to the energy incident on the body at a fixed wavelength and is defined by the following relation:

$$R(x, y, \lambda) = \frac{E_{reflected}(x, y, \lambda)}{E_{received}(x, y, \lambda)} \tag{2}$$

where (x, y) are the coordinates of the pixel, λ is the wavelength, $R(x, y, \lambda)$ is the reflectance, $E_{reflected}(x, y, \lambda)$ is the energy reflected and $E_{received}(x, y, \lambda)$ is the energy received.

In the case of an imaging system, different parameters have to be considered to measure the reflectance. According to a study in face recognition using NIR imaging,[12] the raw measurement $I(x, y, \lambda)$ obtained by the imaging system is given by:

$$I(x, y, \lambda) = L(x, y, \lambda) \times S(x, y, \lambda) \times R(x, y, \lambda) + O(x, y, \lambda) \tag{3}$$

where (x, y) are the coordinates of the pixel, λ is the wavelength, $L(x, y, \lambda)$ is the illumination, $S(x, y, \lambda)$ is the system spectral response, $R(x, y, \lambda)$ is the reflectance of the viewed surface and $O(x, y, \lambda)$ is an offset. Notice that for a given experiment with fixed acquisition parameters and stabilized illumination, $L(x, y, \lambda) \times S(x, y, \lambda)$ and $O(x, y, \lambda)$ can be considered as constants for a given λ. With this hypothesis, by imaging uniform reflective surfaces with a known reflectance values $R(\lambda)$ constant for each pair (x, y), we can calculate $L(x, y, \lambda) \times S(x, y, \lambda)$ and $O(x, y, \lambda)$. Let $I_i(x, y, \lambda)$ the resulting acquisition image of a uniform reflective surface i and $R_i(\lambda)$ its reflectance, we can write (3) as:

$$I_1(x, y, \lambda) = L(x, y, \lambda) \times S(x, y, \lambda) \times R_1(\lambda) + O(x, y, \lambda) \tag{4}$$

$$I_2(x, y, \lambda) = L(x, y, \lambda) \times S(x, y, \lambda) \times R_2(\lambda) + O(x, y, \lambda) \tag{5}$$

By subtracting equation 5 from equation 4 we have:

$$L(x, y, \lambda) \times S(x, y, \lambda) = \frac{I_1(x, y, \lambda) - I_2(x, y, \lambda)}{R_1(\lambda) - R_2(\lambda)} \tag{6}$$

and by replacing $L(x, y, \lambda) \times S(x, y, \lambda)$ in equation 4:

$$O(x, y, \lambda)) = \frac{-I_1(x, y, \lambda) \times R_2(\lambda) + I_2(x, y, \lambda) \times R_1(\lambda)}{R_1(\lambda) - R_2(\lambda)} \tag{7}$$

By replacing $L(x, y, \lambda) \times S(x, y, \lambda)$ and $O(x, y, \lambda)$ in equation 3, we obtain the reflectance formula for our calibrated system:

$$R(x, y, \lambda) = \frac{(I(x, y, \lambda) - I_2(x, y, \lambda)) \times R_1(x, y, \lambda) - (I(x, y, \lambda) - I_1(x, y, \lambda)) \times R_2(x, y, \lambda)}{I_1(x, y, \lambda) - I_2(x, y, \lambda)} \tag{8}$$

According to (8), we calculate the reflectance of each picture of the database after calibration of the system using a white lambertian surface as $I_1(x, y, \lambda)$ and a black lambertian surface as $I_2(x, y, \lambda)$ of know reflectance values respectively $R_1(\lambda)$ and $R_2(\lambda)$. For this experiment the reference surfaces are the white and black squares of the MacBeth color chart.

2.3.4. Exposure Parameterization

The exposure time corresponds to the amount of time a material is illuminated. It has to be adjusted in such a way that the reflectance of the skin surface appears to be maximum for each wavelength combination without saturation of the sensor. With our system, we illuminate the target area with a combination of up to six wavelengths, which implies an increase of the total power dissipation of the LEDs. In addition the reflectance of the skin is correlated with physical particularities such as pigmentation or thickness of the skin. In order to correct possible over-exposure problems, we have characterized the relation between exposure time and power dissipation for eight skin tone models related to eight MacBeth chart color patterns. Figure 5 shows a graphic representation of the eight relations.

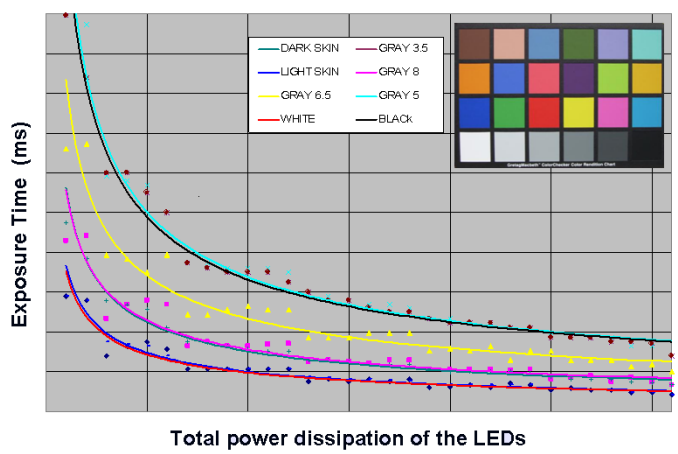


Figure 5. Relation between exposure time and power dissipation for each skin tone model

2.3.5. Acquisition Results

The following pictures (Fig. 6) have been obtained with our NIR acquisition system for different subjects. Each of these images represents one of the 63-image sets prior to any subsequent processing and mapped to the skin tone color reference picture of the subject.

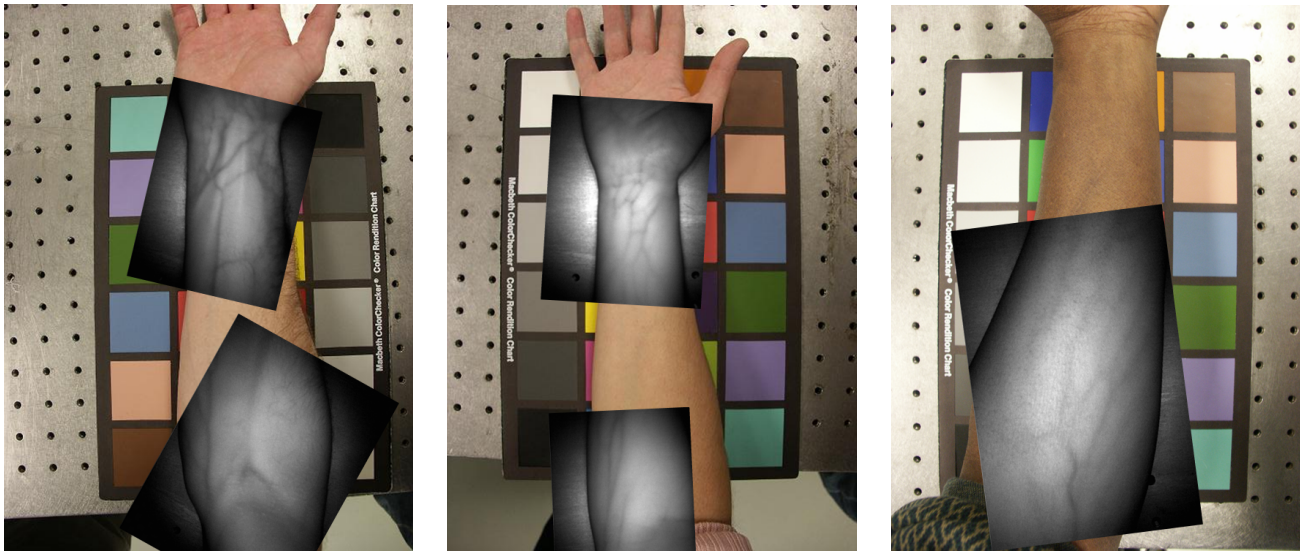


Figure 6. Examples of images obtained with our system for different subjects

3. IMAGE PROCESSING

The next step of our work consists to separate veins from other tissues and to find the path centerline of each vessel. For this experiment, we have simplified our problem by considering a two class system, vein/ not vein, without differentiating other physical particularities. We are processing the data using two approaches designated as supervised learning and semi-supervised learning. Both methods are based on linear dimension reduction approaches: principal component analysis (PCA) and linear discriminant analysis (LDA). With these methods, we want to classify veins and other tissues in two classes by finding the most significant part of their spectral signature. Our methods differ by the process used to create the vein/other tissues mask.
In this section we present and define the algorithms used to process the data and present classification results with our methods. In the last part we present preliminary results on the cross correlation of projection in order to establish the validity of a generalized detection of veins.

3.1. Principal Component Analysis

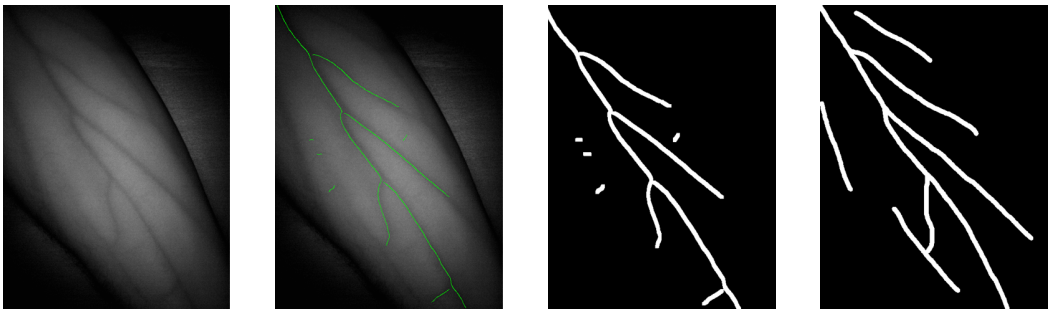
Principal components analysis (PCA)[13, 14] is a dimension reduction technique used to project a M-dimensional space into a lower N-dimensional space in order to simplify the data set. The result of the PCA corresponds to a projection into a new N-dimensional space where the variance of data is maximized on the news axis. In the case of a multi-spectral image, the visual result of the PCA can be compared as a histogram equalization[15] but performed on multiple dimensions.

3.2. Linear Discriminant Analysis

Linear discriminant analysis (LDA)[4] is a dimension reduction technique which maximizes the between class distances and minimizes the within-class scatter so that the class separability can be optimized in the transformed space. Each feature vector of the dataset is labeled using an input classification mask which will be generated manually for the supervised method and automatically for the semi-supervised method.

3.3. Curvilinear Structures Detection

In our NIR images of biological tissues, we can identify veins as a curvilinear structure which appears darker than the surrounding skin due to the absorption-reflection phenomena. Our first approach[1] to detect the veins was to find the transitions between light and dark areas. Using a derivative filter sequence and only local criteria this method tends to be limited in case of a low difference of contrast between vein and skin. According to Steger's road extraction algorithm,[5] another approach for curvilinear structures detection is to consider a line as a function and to identify it according to its differential geometric properties. Our application requires computing the optimal path for the insertion of the catheter i.e. the equation of the centerline of a vein wide enough to receive the catheter. Regarding the previous statement, we have adapted Steger's algorithm for our problem and we present one result of the location of the centerline of the veins (Fig.7(b)), the mask automatically calculated based on the centerline and the approximated width of the veins along this axis (Fig.7(c)) and the mask created manually for comparison with the automatic generation (Fig.7(d)).



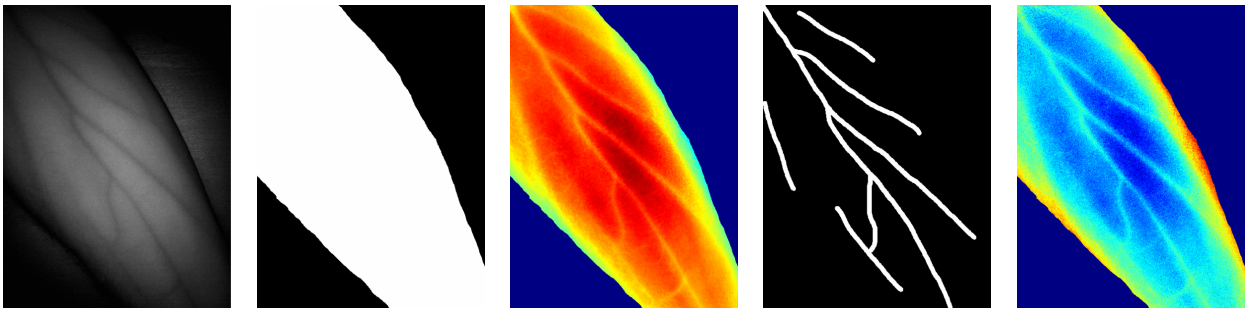
(a) Initial image (b) Veins centerlines detected (c) Mask created automatically (d) Mask created manually

Figure 7. Results of the centerline vein detection process

To generate automatically the mask, we are processing the result of the PCA of the dataset which obviously presents the best contrast vein/not vein regardless the skin tone of the subject.

3.4. Results: Supervised Learning

For this method we define manually the location of the area of interest (Fig. 8(b)) and the location of the veins (Fig. 8(d)). The PCA and LDA are computed only on the feature vectors contained in the region of interest.



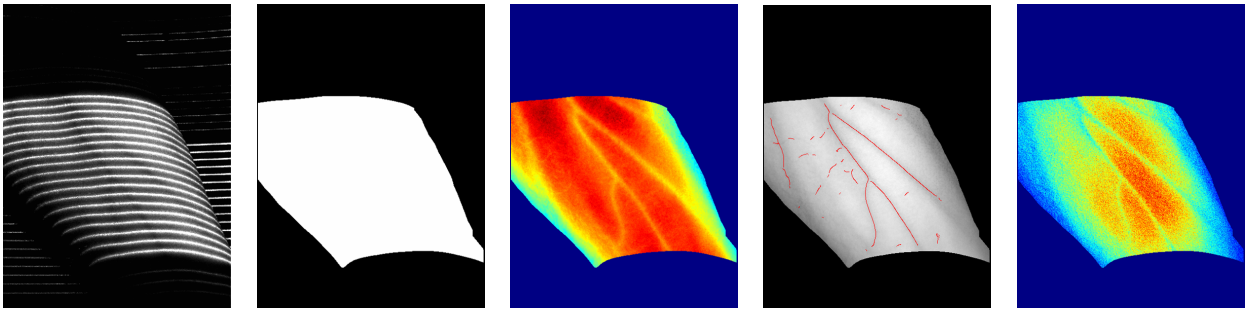
(a) Initial image (b) Mask created manually (c) PCA result (d) Mask veins created manually (e) LDA result

Figure 8. Steps and results of the supervised learning method

The mask vein/not vein (Fig.8(d)) is used to initialise the LDA. Figure 8(c) presents the result of the PCA and figure 8(e) presents the result of the LDA.

3.5. Results: Unsupervised Learning

For this method we define the location of the area of interest (Fig. 9(b)) based on the location of the laser line projection on the surface of the skin. By filling the area delimited by the 19 laser lines (Fig.9(b), we are analyzing the only interesting region in the picture because it is the only one we can use to compute the 3D surface model and so the only one we can use to compute the 3D needle path. In a first step we compute the PCA on the feature vectors included in the ROI (Fig.9(c)) and we use the result to make the vein centerline detection (Fig.9(d)). Finally, after creating the mask for the veins, we are analyzing the area using a LDA (Fig.9(e)).

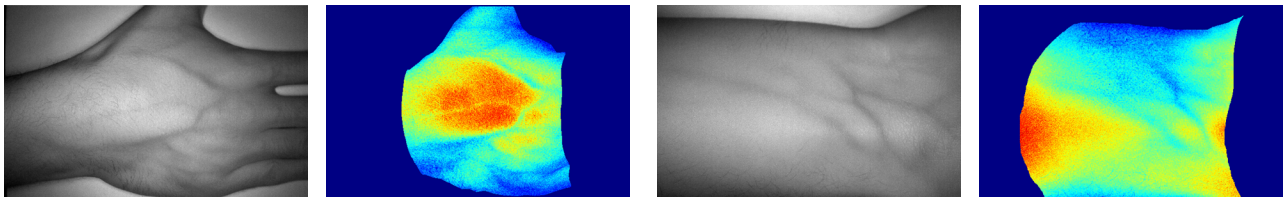


(a) Laser line projection (b) ROI generated from Fig.9(a) (c) Result of the PCA (d) Veins detection on Fig. 9(c) (e) Result of the LDA

Figure 9. Steps and results of the unsupervised learning method

3.6. Cross validation results

Our goal is to find a possible generalization procedure in finding subcutaneous vessels using NIR imaging using the physical characteristics of the skin as only parameters. Each of the 63-dimension vector correspond to the spectral signature of a pixel. Using the LDA approach, we are reducing the dimensionality of each data set to a ten dimension subspace representing a projection of the information in a new definition space which keep the intrinsic properties of the initial data and simplify the separability of the different classes. As a result of the LDA, we obtain a projection matrix which defines the classification parameters of our problem. Considering that each pixel of the dataset has a feature vector corresponding to one of the two classes, our hypothesis was to use the projection matrix calculated for one of the subjects of the database, to project the dataset of another subject and see if we obtain a valid classification result, and if we can see an influence of the skin tone parameters on these results.



(a) Cross validation for a same skin tone than reference (b) Cross validation for a different skin tone than reference

Figure 10. Some results of cross validation

Using the projection matrix computed for one subject (Figure 9(e)), we have projected all the datasets of the database regardless their skin tone properties in a ten dimension subspace. Figure 10(a) represents the results we generally obtain for dataset with a same skin tone and figure 10(b) represents the results we obtain for a dataset with a different skin tone. As you can see, in figure 10(a) the apparent veins have been classified correctly on the center area (with the most homogeneous illumination) and we obtain errors on the offset area where we

have reflection problem due to the topography of the surface of the skin i.e. with the normal to the surface not pointing in the direction of the camera. In figure 10(b), almost all the pixels have been misclassified.

For the moment our database includes only 20 subjects for 100 datasets so we will not conclude about a possible generalization of our method, but the preliminary results tend to point out a correlation between the skin tone characterization and our classification method. In addition for one subject we noticed that two regions of the forearm may need different acquisition parameters to obtain a result which means that we may need to extend our skin model with other parameters such as quality of the skin, body fat mass, etc.

4. CONCLUSION AND FUTURE WORK

Our long-term goal is to develop a fully-automated, vision-guided robotic system for needle insertion and catheterization. In earlier work,[1] we presented a vision-based system to guide a robot that combined the 2D localization of a vein with a 3D model of the surface of the skin. Using this system, we evaluate the hypothesis that different NIR illuminations can be combined to improve the contrast of venous imagery and that the form of this combination may vary under different physiological characteristics. In this paper, we present a system and methods for acquiring and analyzing venous imagery under a variety of NIR illumination wavelengths. We evaluate how various illumination wavelengths can be optimally combined to improve venous contrast. We furthermore examine how the optimal wavelength combinations might vary due to skin tone and/or the presence of hair.

REFERENCES

1. V. Paquit, J. R. Price, R. Seulin, F. Meriaudeau, R. H. Farahi, J. Kenneth W. Tobin, and T. L. Ferrell, "Near-infrared imaging and structured light ranging for automatic catheter insertion," in *Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display*, K. R. Cleary, R. L. Galloway, and Jr., eds., **SPIE 6141**, pp. 61411T–1–61411T–9, SPIE, 2006.
2. T. Vo-Dinh, *Biomedical Photonics Handbook*, CRC Press, Boca Raton, 2003.
3. V. Zharov, S. Ferguson, J. Eidt, P. Howard, L. Fink, and M. Waner, "Infrared imaging of subcutaneous veins," *Lasers in Surgery and Medicine* **34**(1), pp. 56–61, 2004.
4. K. Fukunaga, *Statistical Pattern Recognition*, Morgan Kaufmann, 1990.
5. C. Steger, "Extraction of curved lines from images," in *13th International Conference on Pattern Recognition*, **2**, pp. 251–255, 1996.
6. J. Heikkila and O. Silven, "A Four-step Camera Calibration Procedure with Implicit Image Correction," in *CVPR '97: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, p. 1106, IEEE Computer Society, (Washington, DC, USA), 1997.
7. Z. Zhang, "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations," in *ICCV*, pp. 666–673, 1999.
8. J.-Y. Bouguet, *Camera Calibration Toolbox for Matlab*.
9. P. J. Besl, *Surfaces in range image understanding*, Springer-Verlag New York, Inc., New York, NY, USA, 1988.
10. F. Blais and M. Rioux, "Real-time numerical peak detector," *Signal Process.* **11**(2), pp. 145–155, 1986.
11. J. Forest, J. M. Teixidor, J. Salvi, and E. Cabruja, "A Proposal for Laser Scanners Sub-pixel Accuracy Peak Detector," in *Workshop on European Scientific and Industrial Collaboration*, pp. 525–532, (Mickolj (Hungria)), May 2003.
12. Z. Pan, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25**(12), pp. 1552–1560, 2003.
13. K. Pearson, "On lines and planes of closest fit to points in space," *PhilMag* **2**, pp. 559–572, 1901.
14. H. Hotelling, "Analysis of a complex statistical variable into principal components," *Journal of Educational Psychology* **24**, pp. 417–441, 1933.
15. W. K. Pratt, *Digital Image Processing: PIKS Inside*, John Wiley & Sons, Inc., New York, NY, USA, 2001.