

Exploiting Feature-Based Fusion in LED-based Multi-Spectral Imaging

Herwig Guggi¹, Raimund Leitner² and Bernhard Rinner¹

¹ Institute of Networked and Embedded Systems
Klagenfurt University, Austria
{herwig.guggi,bernhard.rinner}@uni-klu.ac.at

² Carinthian Tech Research (CTR)
Villach, Austria
raimund.leitner@ctr.at

Abstract

Multi-spectral recording systems are used in numerous applications ranging from quality assurance over biometrics to remote sensing. This paper reports on the feasibility of a LED-based multi-spectral imaging system where the spectral characteristics of the illumination is changed by activating different LEDs. The multi-spectral images are captured by a cost-efficient CCD camera. We focus here on applying methods from sensor fusion to increase the classification performance. Various features are generated from the different channels of the multi-spectral images. The most discriminative features are then selected by a forward selection strategy. We demonstrate our feature-based approach in human vein detection. Various test data have been recorded by our prototype of the LED-based multi-spectral capturing system and have served as basis for the experimental evaluation. A detection performance of at least 96 % has been achieved.

1 Introduction

Multi-spectral recording systems are used in numerous applications ranging from quality assurance over biometrics to remote sensing. Multi-spectral imaging (MSI) is a recent technology that allows the acquisition of a spectrum for each pixel (or on images) resulting in superior quality compared to traditional RGB imaging. [7] The objects of interest often have distinctive characteristics at specific wavelengths. MSI exploits these characteristics which typically results in better performance than in monochrome or color (RGB) imaging. MSI systems are also available for UV (ultraviolet), NIR (near-infrared) or even MIR (mid-infrared).

The classical approach to record multi-spectral images is to use spectrographs or tunable filters. Both methods achieve a spectral resolution down to a few nanometers. Gata [4] for example reports that a resolution of 40nm and less can be achieved. However, these capturing systems are one order of magnitude more expensive than monochrome or color cameras. For a number of applications a lower spectral resolution is sufficient which imposes lower requirements on the capturing equipment. One example for such a multi-spectral imaging equipment is a

LED-based capturing system. These systems are comprised of a sensor (and optics) capable to capture images within the overall spectral range of interest. The illumination of the object can be changed by switching on and off LEDs emitting light at different frequencies. Thus, the same optics and sensor are used to capture images of the object with changing but controlled illumination. This setting results in a significantly reduced hardware costs compared to spectrographs or equipment based on tunable filters.

Even though the components are readily available, to the best of our knowledge such LED-based MSI systems are currently not (yet) commercially available. However, the simple hardware setting induces also some disadvantages. First, the spectral resolution is determined by the set of different illuminations which can be generated by the LEDs. Second, since LEDs mounted on different positions are switched on in different patterns, the illumination may become inhomogeneous among the different spectral ranges. The effects of these drawbacks must be compensated in this low-cost LED-based system. Sensor fusion is a mechanism that allows to combine different sensor data resulting in a better result than the source data.

The main goal of this work is to demonstrate the advantages of sensor fusion of such a LED-based multi-spectral capturing system. We focus here on applying methods from sensor fusion to increase the classification performance. Various features are generated from the different channels of the multi-spectral images. The most discriminative features are then selected by a forward selection strategy. The benefit of this feature-based fusion approach is that the features might be easier interpreted by humans and indiscriminate parts of the large multi-spectral space can be discarded from training and classification. This helps to reduce the computational load and memory requirement. We demonstrate our feature-based approach in human vein detection. Various test data have been recorded by our prototype LED-based multi-spectral capturing system and have served as basis for our experimental evaluation. A detection performance of at least 96 % has been achieved.

The remainder of this paper is organized as follows. Section 2 briefly summarizes related work on multi-spectral imaging in medical applications. Section 3 introduces our image acquisition system. Section 4 focuses on our feature-based fusion approach and describes the implemented methods for feature extraction, feature selection and classification in more detail. Section 5 presents the test data used for our evaluation and the achieved experimental results. Section 6 concludes the paper with a short discussion and an outlook on future work.

2 Related Work

MSI has been successfully applied in medical imaging over the last years. The following paragraph briefly describes a few examples:

Wieringa et al. [11] use the fact that in near infrared (NIR) arteries show higher contrast against the myocardium. They recorded images in the spectrum of the visible light and the NIR. The recorded images are then fused to detect (and highlight) arteries. Yamaguchi et al. [12] demonstrate that MSI can be used in dermatology to improve the color reproduction accuracy of skin lesions. The goal of this work was to explore the spectral feature of skin disease using the multi-spectral color enhancement technique and to support quantitative diagnosis. Angelopoulou et al. [1] use MSI in the visible range of the electromagnetic spectrum to determine the skin re-

flectance curve. A pattern which is in 95% directly tied to human skin composition could be figured out. Sets of basic functions have been investigated which enable the development of a model that closely approximates the skin spectral distribution.

The ability to detect vein patterns has applications in medicine and biometrics. Lingyu and Lingham [6] use near- and far-infrared imaging to detect vein patterns for a biometric application. Vein pattern images were acquired from the back of the hand, the palm and the wrist.

Crisan et al. [2] describe a low-cost vein detection system using a LED-based NIR illumination. Their detection system exploits gray-level images at 740nm and applies low-level image processing such as edge detection and thresholding.

Paquit et al. [9] present a system for localizing near-surface veins using NIR imaging and structured light. Their system employs an array of LEDs comprising six wavelengths bands with center wavelengths ranging from 740nm to 910nm. To maximize the vein contrast they determine the optimal combination of bands for a given subject using linear discriminant analysis. However, fusion of the multi-spectral images is not performed. In [8] segmentation is improved using different multi-spectral projection techniques, a broadband illumination source, and accounting for the topography of the skin surface.

3 LED-based Multi-spectral Image Acquisition

Figure 1 shows our LED-based multi-spectral capturing system. Images are captured using a conventional color CCD camera (uEye UI-2210SE) with the NIR blocking filter removed. The scene is lit in turn by 14 LED bands with center wavelengths ranging from 395nm to 940nm. Each band consists of two LEDs: these are arranged into two banks, one on either side of the camera. The LED banks make an angle of 30° to the camera axis.

The response of the camera to different bands varied widely, for this reason it was necessary to adjust the exposure time accordingly. During image capture the exposure time for each band was adjusted to maximize response whilst ensuring that no more than 1% of pixels were saturated. Flatfielding was used to remove any spatial variation in incident intensity.

Figure 2 shows sample spectral images captured with our LED-based image acquisition prototype. Channels 1 to 11 represent images captured with different LED-based illuminations at wavelengths ranging from 395nm to 940nm. Channels 12 to 14 correspond to the native RGB channels of the sensor when the object is illuminated with white light. Channel 15 is the color image of the native RGB channels.

4 Feature-based Fusion

We apply our multi-spectral image acquisition system to the automatic detection of human veins. We treat this as a two-class classification problem. The key steps of this approach are feature extraction, feature selection and classification. First, features are extracted from the multi-spectral images. Second, a subset of distinctive features is chosen. Finally, the selected feature set is used by a classifier. This paper is concerned with image fusion: we will concentrate on steps one and two.

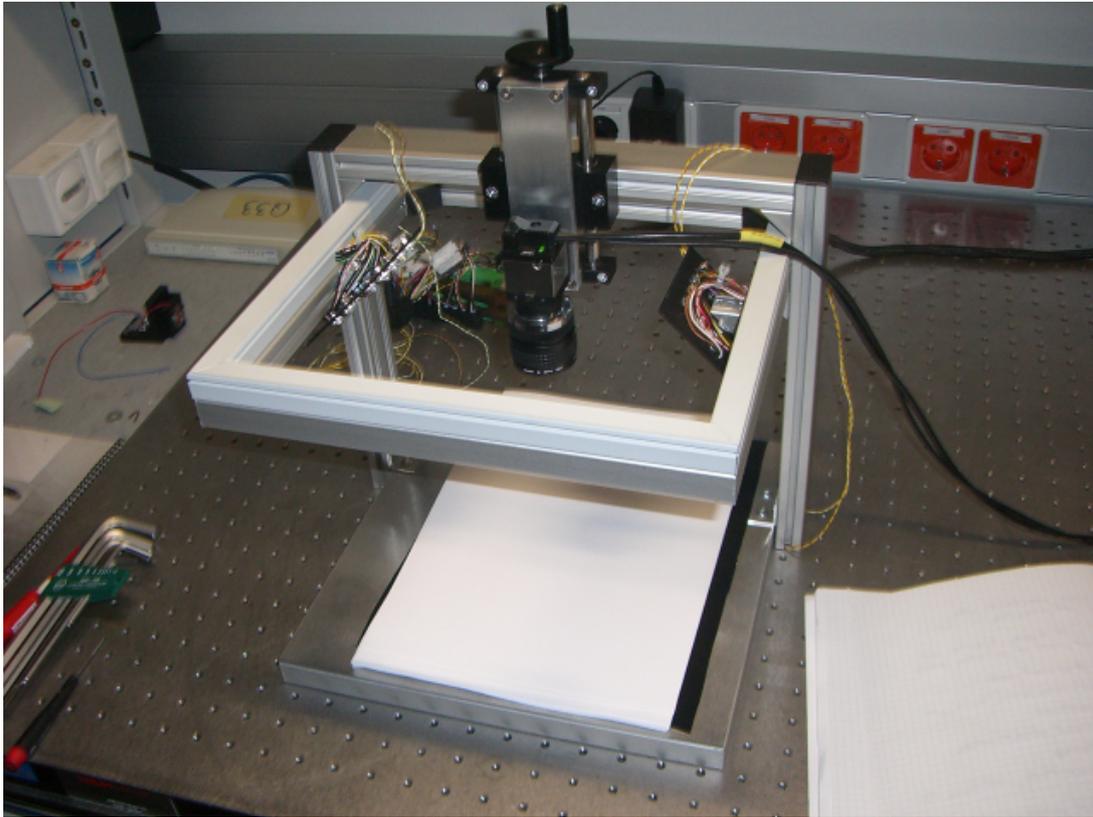


Figure 1: The prototype of our LED-based multi-spectral capturing system.

4.1 Feature Extraction

Features should describe the relevant properties of the data set in a more compact, distinctive and robust form. We focus on the generation of simple features that use the characteristics of a very small spatial neighborhood, i.e., at the level of single pixels or small blocks of pixels. In a pre-processing step, the pixel values of all channels are mapped to $[0 \dots 1]$ by a linear transformation.

Table 1 summarizes the 144 features that were generated from the 14 different channels of the multi-spectral images. These features have been created in a way that the contrast between foreground and background (in our case veins and skin) is improved. The feature extraction methods are categorized as follows. *Basic pixel statistics* such as mean, standard deviation (stdev), median, minimum (min) and maximum (max). *Spatial filters* such as band pass, median and standard deviation filters are applied for generating some features. The block size for these filters affects the quality of the derived feature. A block size of 11×11 pixels achieved a good compromise between suppressing small structures such as hairs and strengthening the larger vein structures. Band pass filters help to reduce the detrimental effects of image noise and hairs. Noise could be successfully attenuated and the contrast between skin and veins was increased. However, we were unable to completely remove the effect of hairs.

Principal component analysis (PCA) is applied to the pixel values of all channels to identify the greatest variance of any projection of the data set. The first principle component is included in the feature set. With some datasets higher order PCA components gave a good contrast between

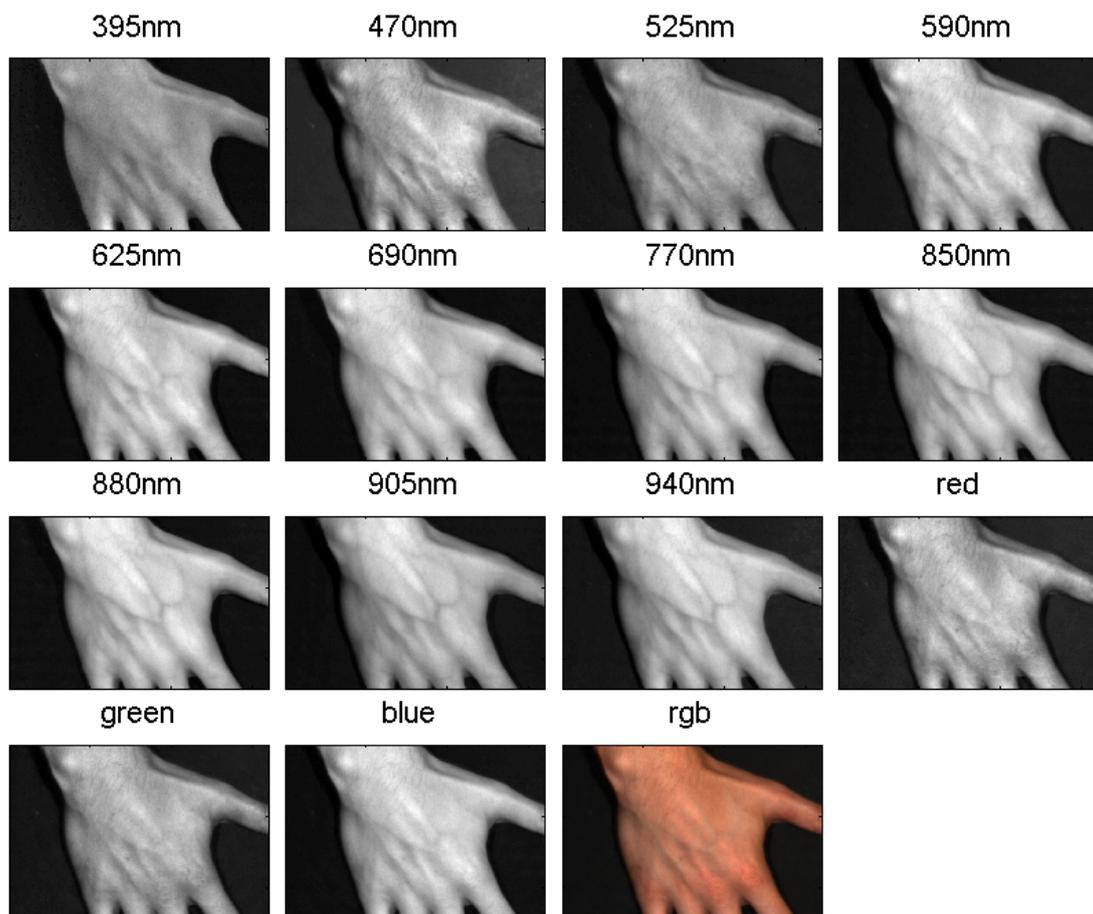


Figure 2: The 15 spectral images (channels) of a human hand captured with our prototype equipment.

| <i>feature index</i> | <i>feature description</i> |
|----------------------|--|
| 1 | first principle component of ch_5 to ch_{11} |
| 2 | stdev-f(f_1) |
| 3 | colorgrad(mean(bp-f(ch_9) to bp-f(ch_{11})), stdev(bp-f(ch_9) to bp-f(ch_{11}))) |
| 4 | mean(ch_2 to ch_4) |
| 5 | mean(ch_8 to ch_{11}) |
| 6 | $f_5 - f_4$ |
| 7 | stdev-f(f_6) |
| 8 | median-f(f_6) |
| 9 | min(bp-f(ch_1) to bp-f(ch_{14})) |
| 10 | max(bp-f(ch_1) to bp-f(ch_{14})) |
| 11 | mean(bp-f(ch_1) to bp-f(ch_{14})) |
| 12. . .25 | bp-f(ch_1) to bp-f(ch_{14}) |
| 26. . .39 | stdev-f(f_{12}) to stdev-f(f_{25}) |
| 40. . .53 | f_{12} - mean-f(f_{12}) to f_{25} - mean-f(f_{25}) |
| 54. . .144 | all pairwise differences of f_{12} to f_{25} , i.e., $f_{12} - f_{13}$ to $f_{24} - f_{25}$ |

Table 1: List of all generated features f_i from the multi-spectral images (channel ch_1 to ch_{14}). These features correspond to the individual pixels of the registered images and are derived using simple operations such as pixel statistics (mean, stdev, min, max) and spatial filters, i.e., band pass filters (bp-f), median filters (median-f) and standard deviation filters (stdev-f), with a block size of 11×11 pixels.

foreground and background. This effect was not true for all datasets which means that they have been identified as less robust for this application.

4.2 Feature Selection

Once the features have been generated a subset of discriminating features must be selected. There are many feature selection methods known in literature (e.g., [5]). In this work we use forward feature selection. This method resulted in a good compromise between feature size, computational load of the selection process and performance of the classifier.

Forward feature selection starts with an empty set of selected features. At each iteration, it adds the *best* feature from the set of unselected features to the set of selected features. The selection process is stopped when adding a new feature does not result in a performance improvement. The two important parameters for forward feature selection are (i) the performance criterion for the feature set and (ii) the selection criterion for choosing the best new feature. We use the performance of the classifier as first criterion which is in our case the *area under the ROC curve (AUC)* [3]. Thus, the AUC is computed for all data sets at each iteration. The result of this computation is stored in a two-dimensional matrix, where one dimension corresponds to the number of features the other to the size of the test data. The second criterion defines what feature to add to the set of selected features. Since the quality of each feature is represented by a vector (one entry for each data set), different criteria are possible.

We compared three selection criteria: (i) minimize the performance variance, (ii) maximize the average performance and (iii) maximize the minimum performance. The first criterion selects the feature which provides the smallest variation among all data sets. The second criterion

selects the feature with the largest performance gain on average, whereas the third criterion focuses on improving the performance of the "worst" data set.

4.3 Classification

The overall goal is to produce a binary image showing the two detected classes. In our case these classes are the veins and the skin. In an initial step good features are identified and an algorithm is trained to fuse those features. The whole process of pattern detection is divided into several phases (image recording, feature extracted, fusion and classification).

At the beginning the images are recorded at different illumination wavelengths. All images are then registered resulting in the multi-spectral image of the object. From these images the features which have been identified in the initial step are extracted. These features are fused with the algorithm (which has been trained in the initial phase). In our case the *linear discriminant analysis (LDA)* is used for feature fusion. LDA is a very well-known technique for 2-set classification and achieves a reasonable performance requiring few computation and memory resources [10]. The result of this fusion process is a monochrome image. Our system uses a global thresholding algorithm to classify the pixels.

5 Experimental Results

5.1 Multi-spectral Data Sets

Our multi-spectral image acquisition system generates a matrix of $14 \times 480 \times 640$ pixels. The 14 channels of this multi-spectral image are registered; each pixel is represented by a scalar value. Seven multi-spectral images have been selected as evaluation data sets for this work. These images have been captured from hands and lower arms of three different persons. The data sets have been labeled by hand to identify pixels representing veins and skin, respectively (cp. Figure 3). The regions of interest are selected in a way that no border regions of the hands are used. For evaluation the data inside these defined regions of interest are used.

For the evaluation, the data sets are partitioned into three groups. The training data is used to train the LDA classifier. The validation data is used to compute the classification performance during feature selection. The test data is used to compute the classification performance with the final feature set. In our evaluation, we partitioned the data set into 15% training data, 15% validation data and 70% test data. For all multi-spectral images, the pixels have been randomly and uniquely assigned to the three data set groups.

5.2 Evaluation of the Feature Selection

Figure 4 compares the different selection criteria. The criterion "minimize the performance variance" (top left) attempts to generate a robust feature set by selecting the feature with smallest performance variations among all data sets. The achieved classification performance is quite good (AUC values between 0.98 and 1.0), however the size of the feature set is pretty large (28 features). The second criterion (top right) selects the feature which maximizes the average classification performance over all data sets. The achieved classification performance is comparable to the first criterion, but the size of the feature set is smaller (23 features). The third

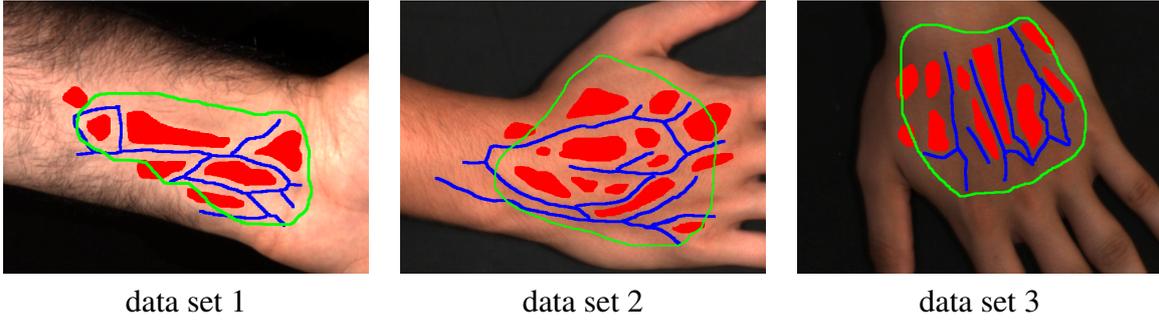


Figure 3: Labeled RGB image of data set 1 to 3 (veins - blue, skin - red, ROI - green).

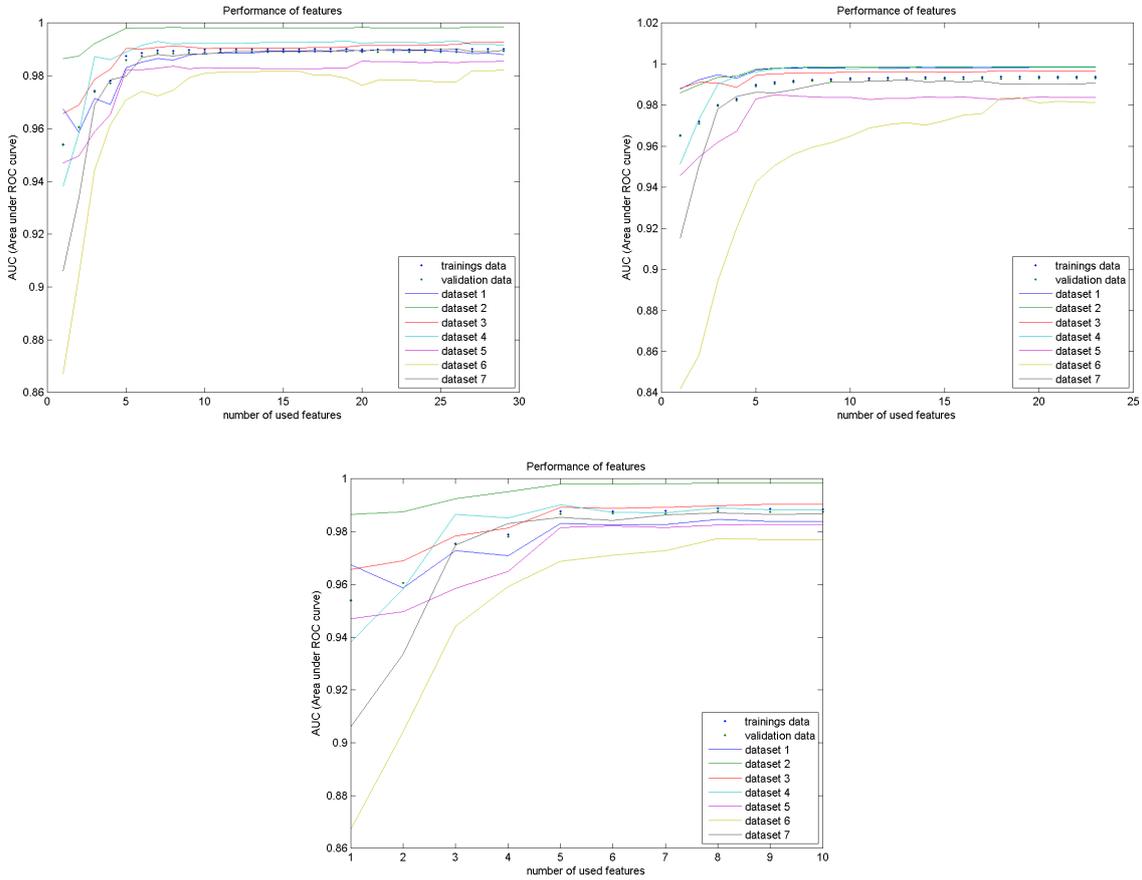


Figure 4: Comparison of the different selection criteria: (top left) minimize the performance variance, (top right) maximize the average performance and (bottom) maximize the minimum performance.

criterion achieves a slightly worse classification performance, however only 10 features have been selected.

When comparing the individual features which have been selected by the three criteria, features f_2 and f_3 are included in all feature sets. Also note that first single feature selected results in a classification performance of 0.94 on average and 0.84 for the "worst" data set.

5.3 Cross Validation

For testing the cross validation the datasets are split to two different clusters. One cluster is used for feature selection and training of the algorithm. The other cluster is used as test data only.

The data that are used for cross validation are seven datasets showing hands and lower arms of three different persons. As the scenes and surface structure of all the test persons differ, the algorithm has to be very robust to provide a good classification for all datasets. For each dataset the region of interest was selected individually. The focus of the work was to separate skinregions and veins. Therefore the regions of interest have been selected in a way that no bordering effects (effects due to bad illumination at the transmission between hand and background) occur.

Figure 5 shows the achieved result. For the feature selection metric maximization of the minimum was used. The horizontal axis shows the number of used features and the vertical axis shows the AUC. The dotted lines represent the datasets which have been used for feature selection and training. The datasets which are displayed with a solid line are only used for testing. It can be seen that all of the datasets which have not been used for training are classified with a quite high performance. The worst performance is achieved using dataset 5, but even here the AUC is larger than 0.96.

6 Conclusion

In this paper we have presented a feature-based fusion approach which has been demonstrated in detecting human veins. The multi-spectral images have been captured by a LED-based capturing system. This prototype captures spectral images with 11 different LED-generated illuminations at wavelengths ranging from 395nm to 940nm. We extract a total of 144 features from these multi-spectral images. The most discriminant features are selected by a forward selection strategy. We evaluated three different selection criteria (i) minimize the performance variance, (ii) maximize the average performance and (iii) maximize the minimum performance. Depending on the selection criterion 10 to 28 features have been selected for classification which based on a standard LDA algorithm.

Seven multi-spectral images have been selected as evaluation data set for this work. These images have been captured from hands and lower arms of three different persons. After training a detection performance of always better than 96 % has been achieved. This promising result leverages further research toward the development of a MSI-based vein-detector on an embedded platform.

There are several directions for future work. First, different and potentially more complex features can be extracted and tested whether they provide more discriminant information. We expect to gain some discriminative power from more spatial features, since the human vein

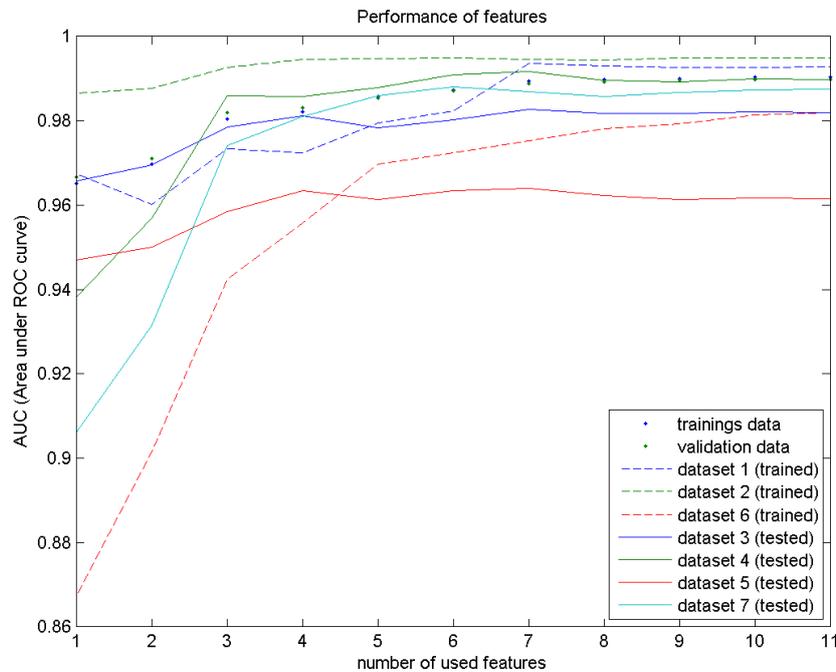


Figure 5: Performance of LDA with maximize minimum metric and cross validation.

patters have a strong spatial characteristic. Second, more comprehensive data recording and experimental tests are required to support our initial performance results. A related future activity would then be to compare our vein detection system with others.

Acknowledgements - Thanks to Gerald McGunnigle for help and suggestions.

References

- [1] Elli Angelopoulou, Rana Molana, and Kostas Daniilidis. Multispectral Skin Color Modeling. Technical report, Stevens Institute of Technology and University of Pennsylvania, 2001. Technical Report MS-CIS-01-22.
- [2] Septimiu Crisan, Joan Gavril Tarnovan, and Titus Eduard Crisan. A Low Cost Vein Detection System Using Near Infrared Radiation. In *Proceedings of the IEEE Sensors Applications Symposium (SAS 2007)*, San Diego, USA, February 2007.
- [3] Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27:861–874, 2006.
- [4] Nahum Gat. Imaging Spectroscopy Using Tunable Filters: A Review. *SPIE - The international Society for Optical Engineering*, Volume 4056:Page 50–64, 04/2000.
- [5] Anil K. Jain, Robert P.W. Duin, and Jianchang Mao. Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 22:4–37, 2000.

- [6] Wang Lingyu and Graham Leedham. Near- and Far- Infrared Imaging for Vein Pattern Biometrics. In *Proceedings of the IEEE International Conference on Video and Signal Based Surveillance (AVSS'06)*, 2006.
- [7] Gianluca Novati, Paolo Pellegrini, and Raimondo Schettini. An affordable multispectral imaging system for the digital museum. *International Journal on Digital Libraries*, Volume 5(Number 3):pages 167–178, 2005.
- [8] Vincent Paquit, Fabrice Mériaudeau, Jeffery R. Price, and Kenneth W. Tobin. Multispectral Imaging for Subcutaneous Structures Classification and Analysis. In *Proceedings of International Topical Meeting on Optical Sensing and Artificial Vision (OSAV'08)*, page 8, St. Petersburg, Russia, 2008.
- [9] Vincent Paquit, Jeffery R. Price, Fabrice Mériaudeau, Kenneth W. Tobin, and Thomas L. Ferrell. Combining near-infrared illuminants to optimize venous imaging. In *Proceedings of the SPIE Medical Imaging*, page 9, San Diego, USA, February 2007.
- [10] Andreas Starzacher and Bernhard Rinner. Evaluating KNN, LDA and QDA Classification for embedded online Feature Fusion. In *Proceedings of the Fourth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, Sydney, Australia, Dec 2008.
- [11] FP Wieringa, F Mastik, DJGM Duncker, AJJC Bogers, C Zeelenberg, and AFW van der Steen. Remote optical stereoscopic multispectral imaging during cardiac surgery. *Computers in Cardiology*, pages 693–696, 2006.
- [12] Masahiro Yamaguchi, Masanori Mitsui, Yuri Murakami, Hiroyuki Fukuda, Nagaaki Ohyama, and Yasuo Kubota. Multispectral color imaging for dermatology: application in inflammatory and immunologic diseases. In *Proceedings of the IS&T/SID 13th Color Imaging Conference*, 2005.