

*Julian Bader, Björn Labitzke
Marcin Grzegorzek, and Andreas Kolb*

Multispectral Pattern Recognition Techniques for Biometrics

Keywords:

*biometrics,
multispectral imaging,
pattern recognition*

1. Introduction

It has been quite a while since the first optics-based biometric identification systems were introduced. Nowadays, there exist commercial imaging techniques in the visual spectrum (VIS) to detect a variety of different features used for classification, such as faces, palmprints or fingerprints. Although the reliability of such systems is quite high, for most techniques it is crucial to acquire the probes under certain controlled conditions (e.g. illumination). In turn, the hardware used to perform a proper authentication is costly and the time the authentication process takes might be unfeasibly high. In addition, one has to make sure that the identification database contains meaningful samples of an appropriate quality.

In order to relax the requirements on the controllability of external conditions, multispectral imaging can be used. Here, information of a feature in several spectral bands is acquired and fused. Therefore, one might get enough information to compensate uncontrolled conditions, such as insufficient illumination or decreased scanning resolution.

Surveying the literature, we realized that the work in the field of multispectral biometrics differs according to the feature under observation. Also, there exist similarities among methods which perform recognition on the same feature. Thus, in this survey, we classify the various approaches according to this observation. Here, we distinguish between recognition of faces, palmprints, fingerprints and iris.

In the next section biometric identification using faces is described. There exist mainly two approaches which differ on the spectra used for recognition. While the one method (e.g. [2] and [21]) considers fusion of broad band visible spectra with thermal infrared, the other (e.g. [5]) acquires several narrow bands and decides which of them carry the most significant information.

Section 3 presents identification methods based on palmprint scans. As various work has been proposed, the approach in [26] is described in more detail. Additionally, the methods described in [9] and [13] are presented whereas only the differences to [26] are explained.

In Section 4, it is dealt with multispectral fingerprint recognition. The work of two papers is presented which perform recognition using comparable acquisition systems.

In Section 5, an initial work on multispectral iris analysis by Boyce et al. [1] will be discussed. Most commercial systems use only the near-IR wavelengths for iris analysis and completely disregard the possibility of employing multispectral information in this application domain. We will show in this section that using multispectral data for iris-based human identification might be very beneficial when applying suitable pattern recognition techniques.

The last section concludes this chapter. For all presented methods, the similarities and differences are stated. At the end, an outlook for future research is given.

2. Face Recognition

For the VIS range, image-based face recognition approaches perform quite well on probes which have been acquired under controlled conditions. Usually, a picture (monochrome or RGB) of the face is taken in frontal view under uniform illumination. In order to maximize the performance, people have to express a neutral facial pose. In a pre-processing step the actual face is segmented (and detected) and the picture is normalized. Afterwards, the significant features are extracted (e.g. with the eigenfaces approach [18]). For the actual classification, the extracted features are compared to a previously acquired gallery of known subjects. In order to decide to which gallery image the probe belongs to, the similarity score is computed according to these features.

However, for a variety of real world scenarios, it is not possible to acquire images of a certain quality and with controlled view, illumination and facial expression. In the majority of cases, it is necessary to compensate such flaws on poorly recorded images which do not comply with the above mentioned criteria.

In the literature, there exist mainly two ways to use multispectral imaging to improve the recognition rate. The first one fuses broadband infrared images with monochrome images in the visible spectrum. Here, the idea is to use the advantages of both spectra. The passive infrared image acquisition can compensate insufficient illumination but has problems to accurately render glasses what, in turn, can be compensated in the VIS spectrum.

The second approach also focuses on balancing out bad illumination conditions. In contrast, it uses several narrow bands in the visible spectrum. These bands are selected according to how good they perform in face recognition when only this one band is used. By taking into account the optical properties of the measurement system and light source, the images gathered at different wavelengths are normalized.

2.1. Broadband Visible and Infrared

The approaches proposed in [2], [10], [19], [20] and [21] use two facial images which have been acquired using two different broadband spectra, visible and infrared. For the infrared spectrum images at different wavelengths have been taken. Namely, the Long Wave or thermal IR (8 - 12 μ m), the Mid Wave IR (3 - 5 μ m) and the Near IR (900 - 1700nm) have been used. The images in the visible and LWIR domain are coregistered. All pictures have been taken from the Equinox face dataset¹.

According to [21], data fusion can take place at different stages in the pattern recognition pipeline. In [3], [10] and [19], the authors describe methods to fuse visible and infrared data at the image level and feature level. The authors of [21] proposed a method which combines fusion at both levels. For all latter discussed approaches, it is assumed that necessary preprocessing steps (e.g. alignment of faces, segmentation of the face area) already have been done.

According to [19], thermal infrared images of the human face have a comparably low resolution and therefore contain lower frequencies. In contrast, images in the visible spectrum have a higher resolution and also capture small details which results in higher frequencies. In order to take into account these circumstances, the image-based fusion is done in the wavelet domain. Here, an image is decomposed in its various frequency components. For in-depths information about wavelets, the reader is referred to [7]. However, the result of this wavelet transform is a set of coefficients (similar to the Fourier transform) that describe the local occurrence of frequencies in the image.

¹ see <http://www.equinoxsensors.com/products/HID.html>

The goal is to choose those coefficients which carry the most significant information regarding face recognition. Thus, a new fused image can be reconstructed that is processed further in the pattern recognition pipeline.

In order to find a good (however not guaranteed optimal) choice of wavelet coefficients, in [3], [10] and [19] a Genetic Algorithm (GA) is employed. For more detailed information about GAs, the reader is referred to [8]. Here, only a short overview is given. GAs try to find the best solution for a problem iteratively. Possible solutions are encoded as strings which represent chromosomes. A set of chromosomes is called a population. In each iteration three things can happen to the population:

- Selection: The n best solutions for the problem are kept. The others are discarded. To evaluate the quality of a solution, a fitness function is introduced.
- Crossover: The chromosomes of two solutions are combined. Thus, a new solution is created which contains parts of its two parents.
- Mutation: One or more characters of a string are changed randomly. Thus, new solutions are created which are not part of the original genetic material.

If a solution of the population reaches a certain quality (e.g. 100% recognition rate of the training set), the algorithm converges.

For the image fusion problem, the chromosome is a bit string with the length l resembling the number of wavelet coefficients. For each coefficient the corresponding bit states whether to take the wavelet coefficient from the IR or from the VIS domain. The initial population starts with n chromosomes. Here, the chromosomes are initialised with zeros. Then, the number of ones in the string is chosen randomly and the ones are distributed at random. After each iteration, the number of chromosomes doubles (due to crossover) to $2n$ from which the best n solutions are selected. In order to determine the fitness of a solution, the images are fused according to the bit string. Afterwards, face recognition is performed and evaluated using a validation dataset. For the crossover strategy, uniform crossover is used. In their experiments, the authors chose a crossover probability of 0.95. Mutation flips one bit of a chromosome with the probability of 0.03.

The feature-based fusion method described in [3] and [19] also employs GAs which work in a similar manner as for the image-based fusion. For both

images (VIS and IR), the features are extracted using the eigenface approach that is based on principle component analysis (PCA). Further details regarding eigenfaces can be found in [18] and [23]. After the feature extraction, again a Genetic Algorithm is employed to determine the most significant features of the IR and visible domain. Here, for each spectrum the first 100 eigenvectors are considered. Again, a bit string (of length 100) represents the chromosome and each bit determines whether to use the feature from IR or the visible band. The rest of the algorithm works exactly in the same way as the image-based fusion.

In [21] an approach has been proposed which combines both fusion methods. Here, the authors abandoned the use of Genetic Algorithms. As they say, GAs are too expensive to compute. However, the image-based fusion still depends on wavelets whereas a deterministic fusion method is proposed. This method averages the coefficients in the approximation band of both spectra. For the detailed bands, decision maps are computed that determine how they are fused. In addition, the feature-based fusion differs from the method mentioned above. After face detection and segmentation, the resulting image is converted into polar coordinate form and transformed into the 2D Fourier space. This Fourier transform of the face is convolved with the 2D Fourier representation of the log polar Gabor Wavelet. The Inverse Fourier Transform of this convolved image results in a matrix containing complex values. These values correspond to amplitude and phase information of the fused facial image.

A dual ν -Support Vector Machine (2ν -SVM) is applied to decide whether to use amplitude or phase information. For this purpose, both features are grouped in windows of 3 by 3 pixels. For each window one of the two features is chosen (further details can be read in [21]). For classification, the amplitude and phase information of the fused image are used. Here, these features of the two images that should be compared are divided into a number of frames. For two corresponding frames, the distance between each of the two features is determined by a distance metric. This distance is then used for score computation. If this score exceeds a previously defined threshold, it is considered as a match.

In [20], the same authors proposed another two-level fusion approach. Instead of the feature-based fusion at the second level, score-based fusion is used.

2.2. Narrowband Visible

In contrast to the approach described above, the dataset used in [4], [5] and [6] consists of narrow spectral bands in the visible domain. For acquisition, a liquid-crystal tunable filter (LCTF) mounted on a monochrome camera is used. For one person, images have been required in the VIS range with 10nm steps. Thus, for each subject, a variety of images have been acquired. Also, every person has been recorded multiple times under varying illumination conditions where the spectral power density of each light source has been determined by a spectrometer.

The focus of this approach is mainly to compensate misclassifications resulting from images acquired under different illumination conditions. In order to achieve this, the optical properties of the whole measurement system are considered. Also, the spectral power density (SPD) of the light source is taken into account.

Neglecting the scattering of light on curved surfaces of the face, the intensity of a pixel p measured by the monochrome camera can be described as follows

$$p = \int_{\lambda_{\min}}^{\lambda_{\max}} R(\lambda)L(\lambda)S(\lambda)d\lambda, \quad (1)$$

where R is the reflectance of the object, L is the luminance resulting from the SPD, S is the spectral response of the camera and λ is the wavelength. Here, it is assumed that the light is uniformly distributed throughout the scene. An LCTF is mounted on the camera. If a certain spectral band with the peak at wavelength λ_i is selected, light between $\lambda_{i,\min}$ and $\lambda_{i,\max}$ is transmitted. This results in the pixel intensity

$$p_i = \int_{\lambda_{i,\min}}^{\lambda_{i,\max}} R(\lambda)L(\lambda)S(\lambda)T_{\lambda,i}(\lambda)d\lambda, \quad (2)$$

for a peak wavelength λ_i , where $T_{\lambda,i}$ is the transmittance of the LCTF for λ_i . Assuming that the selected bands are relatively narrow, it is sufficient to take only one sample for each peak wavelength λ_i . Thus, the above formulation can be simplified to

$$p_i = R_{\lambda,i} L_{\lambda,i} S_{\lambda,i} T_{\lambda,i}. \quad (3)$$

The parameters L and S are determined by the measurement system. The parameter R is also considered as constant as long as the same object/subject is recorded. Therefore, the appearance of the image only depends on the SPD of the light source. The intensity of a pixel

$$p_{1,\lambda,i} = F_{\lambda,i} L_1 \quad (4)$$

and a pixel

$$p_{2,\lambda,i} = F_{\lambda,i} L_2 \quad (5)$$

which have been acquired for a certain peak wavelength λ_i only differ by the type of illumination. Here, F stands for the product of R , S and T . As both pixels cover the same area and the properties of the optical acquisition system keep the same, F is equal for both pixels. In order to cancel out the illumination effect on pixel $p_{1,\lambda,i}$ we have to divide the intensity by L_1 . Doing this for all pixels in an image results in a normalized image which is independent of the illumination. If the normalized pixel $p_{\lambda,i}$ is multiplied with L_2 , the intensity of the pixel $p_{2,\lambda,i}$ is achieved. By applying this transformation to all pixels of an image, two images taken under different illumination conditions can be made comparable. However, this transformation only works under the conditions that the scenes have been illuminated uniformly and that the scattering of light on the facial curvature can be neglected.

If we have a closer look at the transmittance T of the LCTF and the spectral response S of the camera (see fig. 1), we can see that the intensity varies depending on the wavelength. This means that some images (especially at shorter wavelengths) are darker than other ones. In turn, when fusing several narrow band images, their contributions are not uniform. In order to compensate this effect, the contribution of each band has to be

biased by a weighting factor $w_{\lambda,i}$. By taking the reciprocal of S and T as bias, the influence of the optical system is cancelled out.

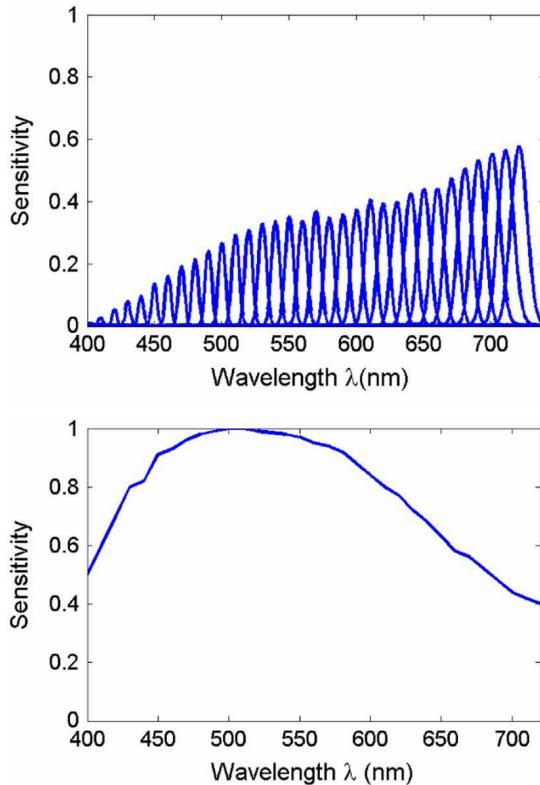


Fig. 1. Transmittance of a liquid-crystal tunable filter (left) and spectral response of a charge-coupled device (right) [5]

A simple fusion strategy is now to sum up the weighted intensities of all acquired narrow bands

$$p = \frac{1}{C} \sum_i w_{\lambda,i} p_{\lambda,i}, \quad (6)$$

where C is the sum of all weights $w_{\lambda,i}$. Thus, one receives a broadband grayscale image with uniform contribution of all measured spectral bands.

The above mentioned fusion strategy assumes that each spectral band carries an equal amount of information. However, this might not be true as the spectral reflectance of skin is not uniformly distributed among the visible spectrum. In fact, higher frequent light is absorbed more than red light.

There are also some characteristic peaks which result from the different chromophores (melanin, carotene and hemoglobin) present in the skin. Thus, the authors of [6] proposed a method to determine the information quality of narrow spectral bands.

To do so, first two sets of multispectral images are acquired. Each set contains samples of N different persons. Each sample consists of images at k narrow spectral bands. It is assumed that all samples have already been preprocessed (e.g. aligned and registered) and the features are already extracted². S_{ij}^k is defined as the similarity score of the images of two persons i and j whereas the images of the k -th band have been compared. For each band, the similarity score can be divided into two sets. The first one is the genuine set which is defined as $G_k : \{S_{ij}^k, i = j\}$. It contains all scores for which the two compared images show the same person. The second one is the imposter set defined as $I_k : \{S_{ij}^k, i \neq j\}$. I_k consists of all similarity scores for which the two compared images show different persons.

Having determined these sets for all spectral bands, the probability density function (PDF) of all sets is computed. In the ideal case the PDF of genuine and imposter set do not overlap. Thus, a person always could be identified correctly. However, in reality these two PDFs do overlap. In order to determine how well a spectral band separates these two sets the Jeffrey divergence $Q_{JD,k}$ is computed. For additional background knowledge, the reader is referred to [6] and [17].

By ranking the $Q_{JD,k}$ values, the importance of each band for face recognition becomes clear. When having a certain number of n bands which should be considered, the first n bands according to the rank are taken. If it is not clear how many bands to take, the degradation percentage

$$\eta_k = \frac{Q_{\max} - Q_{JD,k}}{Q_{\max} - Q_{\min}} \quad (7)$$

can give a good hint. Here, Q_{\max} and Q_{\min} are the highest and lowest divergences value. For a previously defined threshold η_{th} , all bands are chosen for which $\eta_k \geq \eta_{th}$ holds.

² Feature extraction and computation of similarity scores have been done by proprietary software.

2.3. Summary

In this section two multispectral face recognition methods have been presented. The first method uses the fusion of broad band visible and infrared images to improve recognition performance. The idea behind is that uncontrolled conditions such as poor illumination or facial expressions can be compensated by the infrared spectrum. The second method considers several narrow bands in the visible spectrum. Here, the primary goal is to normalize the acquired images according to the light source and the optical system. Additionally, the number of spectral bands are reduced by only considering these bands which carry the most significant information.

In order to decide about which features are fused, the authors of [2], [9] and [19] proposed Genetic Algorithms. They state that they do not know how the wavelet coefficients depend on each other. Thus, they use uniform crossover to merge the features of two prior determined solutions. In our opinion, GAs should not be used due to their randomized nature. The quality of the results depends highly on the initialization of the algorithm. Additionally, the unawareness of the dependence of wavelet coefficients severely decreases the performance of the employed GA as the crossover step hardly creates usable results. By understanding the wavelet space, the fusion strategy can be improved. In addition, a more suitable learning algorithm can be employed.

In [4], [5] and [6], the authors use the spectral power distribution of the illumination to normalize the acquired multispectral images. While for some outdoor images it might be possible to actually know the SPD in advance, for the majority of cases with uncontrolled conditions this knowledge is not present. Also, by assuming a uniform illumination throughout the whole scene, the darker parts or parts of the face which do not point directly into the camera are computed wrong. In order to normalize a picture correctly, the position and directional characteristic of the light source has to be known. Additionally, the surface normals of the face have to be considered as the reflectance depends on the angle of incoming and outgoing light.

3. Palmprint Verification

Conventional palm recognition methods (e.g. [11]) use grayscale images for personal identification. In contrast to face recognition, the acquisition of such images always takes place under controlled conditions. Thus, the primary goal of multispectral approaches is to further increase the recognition rate by taking a larger amount of information into account. According to [9], [16] and [26], multispectral imaging for palmprint recognition is also suitable in terms of spoof detection.

In the literature, there exist a variety of different approaches which range from the evaluation of fusion methods (e.g. [12]) to the description of whole systems such as [16] or [26]. The overall principle remains roughly the same: After acquisition, the images are pre-processed (alignment, selecting the region of interest, etc.). Then, the significant features are extracted. These features are used to compute a similarity score between the acquired image and the image gallery. According to these score values, the actual classification is done.

However, at some point in this whole process, the information of the several spectral bands need to be fused. We distinguish between two classes of approaches. The one class performs image-based fusion. Similar to the formerly mentioned face recognition approaches, the images are merged prior to the feature extraction. The second class uses score-based fusion. Here, for each spectral band, a similarity score is computed. These scores are accumulated (e.g. by using a *majority vote* technique) to make the final decision.

In the rest of this section, one of the palmprint recognition approaches is described in detail to illustrate the use of multispectral imaging in this field. Afterwards, differences to other methods are discussed.

3.1. A score-based fusion approach

In [26], Zhang et. al. described a complete multispectral palmprint recognition system which uses score-based fusion. One of the goals of the authors was to take into account the correlation of information between the spectral bands. Thus, one can find a minimal combination of bands which yields satisfying recognition results.

The acquisition equipment consists of a box with a pane of glass on top. Inside the box, at the bottom, there is a CCD that takes pictures of a hand placed on the glass. The box also contains four arrays of LEDs with their peak wavelength at 470nm, 525nm, 660nm and 880 which correspond to blue, green, red and near infrared light respectively. When acquiring a dataset, four images are captured each with one array of LEDs turned on while all the others are turned off. Thus, each image represents one spectral band.

As the hand is held still during the acquisition process, no registration needs to take place. Only, cropping to a rectangular region of interest is done in order to reduce the amount of input data for feature extraction.

For feature extraction, a texture-based coding approach, namely Gabor filtering, is used. A Gabor filter tries to find edges according to a user defined orientation. In this approach six Gabor filters are applied for different orientations. As the characteristic lines on a palm are darker than the surrounding skin, for each pixel, the orientation with the least response is taken into account as feature. The result of this operation is a feature image where the prevalent orientation of each pixel is coded by three bits. The bit values for each orientation are chosen in such a way that the Hamming Distance³ of the bit string represents the distance between two orientations.

Using this per pixel orientation coding scheme, it is easy to determine the similarity between two palm images:

$$d(P, Q) = \frac{\sum_{x=0}^M \sum_{y=0}^N \sum_{i=1}^3 (P_i(x, y) \otimes Q_i(x, y))}{3MN} \quad (8)$$

Here, P and Q correspond to the feature images of two palm images. i refers to the i th bit of the feature image where the bit string of the pixel at position (x, y) is considered. M and N are the dimensions of the images. In order to take misalignment into account, one of the images is shifted in the range of -3 to 3 pixels and the smallest value for d is regarded as the final result.

According to the authors, it is not feasible to simply accumulate the score values for each band to determine an overall score. Since certain

³ The Hamming Distance is defined as the number of differences of two strings.

information is correlated among the several spectral bands, by summing up the score, this information would be counted several times and thus have a higher impact. The authors state that principal lines on a palm are present in all bands. By simply summing up the score values of different palms with the same principal lines, they would be spuriously matched. Thus, the following equation is employed which accounts for information which is equal among two bands

$$d_{F_1 \cup F_2} = d(F_1) + d(F_2) - d(F_1 \cap F_2) \quad (9)$$

F_1 and F_2 are tuples (e.g. (P, Q)) containing two feature images of two different palm images. These palm images have been acquired at the wavelengths λ_1 and λ_2 . The intersection of these two sets can be seen as the fractional part of information which is overlapping for both spectral bands. This can be computed as follows

$$d(F_1 \cap F_2) = \frac{d(F_1^i, F_1^j) + d(F_2^i, F_2^j)}{2} P_{OP}(F_1, F_2) \quad (10)$$

The superscripts i and j refer to the two samples which are supposed to be compared. P_{OP} is the percentage of overlapping information between two wavelengths λ_1 and λ_2 . While the similarity scores for two feature maps can be computed easily, the percentage of common information needs to be determined empirically. The authors of [26] did this by taking the mean of the comparison of 250 samples for each band combination. Here, the standard deviation was between 4% and 5.5%.

3.2. Feature band selection

In [9], another approach was proposed that aims at selecting a minimal set of spectral bands. However, while the approach mentioned above uses RGB and near infrared, for this method, 69 narrow spectral bands in the range between 420nm and 1100nm are considered. The spectral images are acquired using an LCTF. As it is assumed that the hand does not move during acquisition, the region of interest is selected by cropping each image. Here, the (2D)2PCA method [25] is used for feature extraction. As the name

suggests, this method is based on principal component analysis. Two covariance matrices are created where one is computed along the columns and the other along the rows. For each covariance matrix, a matrix of eigenvectors is created. A threshold determines how many eigenvectors are used to transform the gallery and the test image. After transformation, the gallery image is subtracted and the norm of the resulting matrix is computed and used as score value.

The images which produce the smallest score are to be considered as a match. Score level fusion is now employed by summing up the values for each band in consideration. However, the score values are normalized using the mean and standard deviation according to the training set. In order to determine the most significant feature bands, the recognition rates for each feature band and for each combination of two and three bands are evaluated.

3.3. Image-based fusion approach

In [12], Hao et. al. evaluated different image-based fusion methods. One year later, in [13], the authors describe a whole biometric identification system.

This system uses a setup quite similar to [26]. However, the hand is not placed onto a glass. Instead, it is held still in front of a camera. Also, the system consists of six arrays of different LEDs (rather than four) with bands in the visible and infrared spectra.

Because it cannot be guaranteed that the hand does not move during the acquisition process, as a first step, the images need to be registered. Afterwards, the ROI is cropped.

According to the authors an image fusion scheme includes:

- **multiscale decomposition** which transforms the pixel representation into a more suitable space (like wavelet domain),
- **activity measure** which judges the quality of the input,
- **coefficient combining** which defines how the input is combined according to the activity measure.

In contrast to the fusion method of [3], [10] and [19] described in the face recognition section, no search algorithm is employed which tries to find the optimal set of coefficients among the different spectral bands. Here, the

chosen coefficient combining method decides about coefficients to be used. After the image fusion, the features are extracted and a similarity score is computed.

3.4. Summary

Three different multispectral palmprint recognition methods have been presented in this section. These approaches can be distinguished by the number of spectral bands taken into account. Additionally, they differ in the fusion methods. While some fuse the information at the image level, others accumulate the distance or score values. In addition, the approaches proposed in [9] and [26] are of interest as they consider the correlation between the several spectral bands. This could be further investigated in order to receive a feature extraction method which regards the multispectral information as a whole.

The topic of spoof detection which was mentioned in [9], [16] and [26] is only addressed on the side or not at all. The authors of [26] were the only ones who performed an anti-spoofing test. They printed out the blue channel on a paper and used it as a sample for their system. As expected, using only the blue channel for verification, results in a match. Acquiring and comparing the green, red and NIR channel additionally rejects the sample. This kind of simple spoofing test does not have a significant meaning. Due to the sample being based on the information of only one channel, it will of course be rejected as soon as other channels are considered. Thus, it would be more interesting how the system behaves if the combined RGB channels are printed out. However, the system is based on an RGB camera which tries to capture exactly those details which are of interest for the human eye. As long as a human observer cannot distinguish between a real and a fake hand, the system also does not.

4. Fingerprint Biometrics

Conventional optical fingerprint sensors are typically based on total internal reflectance (TIR). The results of the acquisition are TIR images, which are usually examined to detect or match minutiae (Galton details [24]). The performance of the conventional TIR sensors is generally degraded by certain common occurrences, such as inconsistent contact

(e.g. due to little pressure or dry skin) [24]. Furthermore, the performance of conventional optical fingerprint sensors also is correlated to the size of the sensor itself that means that the performance degrades as the size decreases [15].

In the following, two approaches are presented, that describe multispectral based solutions to improve the performance of conventional fingerprint sensors. In “*Multispectral Fingerprint Biometrics*” Rowe et al. [14] described a novel fingerprint sensor that combines a multispectral imager and a conventional TIR sensor. The goal is to improve the usability and reliability of security related standard technology, especially in terms of these above mentioned occurrences which prevent proper acquisition. Since optical fingerprint sensors are used in a broad range of applications, like access control or even in consumer devices such as laptops and cell phones, there is an emerging need to reduce the physical size of the sensors in order to reduce the area of the device that is occupied by the sensor. Therefore, in “*Biometrics Based on Multispectral Skin Texture*” Rowe [15] examined the feasibility of MSI based small-area sensors and compared the performance to traditional TIR sensors. Before these papers are presented in more details, the concept of the underlying MSI based sensor will be described in a short overview.

4.1. Concept of MSI based fingerprint sensors

Both approaches are using a conceptually comparable sensory system, which is described in the following. There are only a few small differences (e.g. different wavelengths) between both systems, which for now are ignored and are discussed in the associated sections later on.

Figure 1 shows a conceptually illustration of the major optical components of the suggested MSI sensor. Multispectral illumination is realized by several LEDs that are packaged to different groups of wavelengths. The different wavelengths penetrate the skin to different depths and are absorbed and scattered differently by various chemical components and structures in the skin [15]. An orthogonal configuration of polarizers is used to emphasize the scattered light by

reducing the influence of directly reflected light from the surface of the finger (cf. fig. 1).

Especially for the second approach, also a non-polarized illumination path is included, which allows the acquisition of images that are much stronger influenced by the surface features of the finger [15]. Additionally to the described multispectral illumination, the sensor also includes conventional TIR components to maintain full TIR imaging functionality.

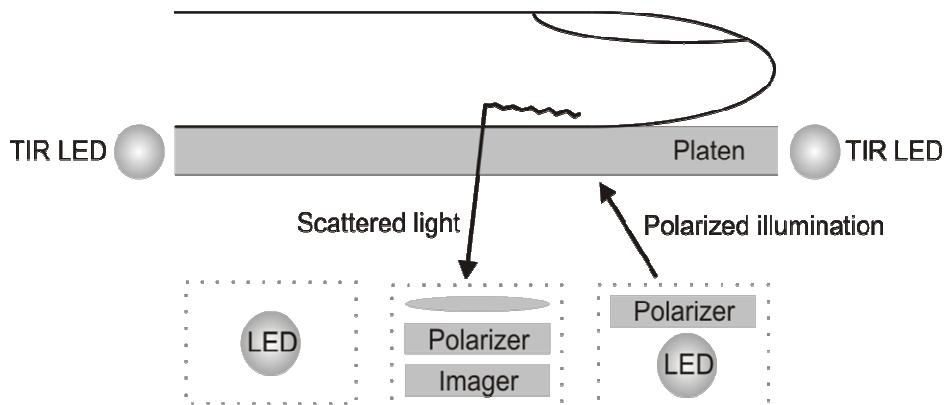


Fig. 2: Major optical components of the suggested MSI sensor.
(Illustration is inspired by [14, 15]).

4.2. Multispectral Fingerprint Biometrics

As mentioned before, Rowe et al. [14] have developed a novel fingerprint sensor that allows the acquisition of both TIR and multispectral images with the goal to improve the performance of fingerprint sensors regarding problems which prevent sole TIR sensors from working. The described concept of the sensor was applied in this approach except the non-polarized part. The polarized multispectral illumination covers six wavelength ranges (400, 445, 500, 574, 610 and 660nm).

In the acquisition process, a series of MSI images and a single TIR image is acquired. Additionally, also two ambient light measurements were taken without any LED illumination before and after the MSI series. During the acquisition of the TIR image by red

(640nm non-polarized) illumination, also a seventh multispectral image is captured.

While TIR images are processed as raw data, the multispectral images are preprocessed with a sequence of steps:

- The two ambient measurements are averaged together and subtracted from each MSI image for ambient-light correction.
- Multispectral images are smoothed with a Gaussian filter and then are used as divisor for the original images to remove the effect of non-uniform illumination across the platen.
- Bandpass filtering is performed to suppress out-of-band variation (especially pixel-to-pixel noise) while important frequencies for fingerprint features are passed.
- Finally, adaptive histogram equalization is done to equalize the fingerprint contrast of an image.

The authors have used commercial fingerprint software⁴ to perform the task of fingerprint detection and matching. In a multi-person study (15 adults, each 4 fingers and 6–8 times) they collected 602 samples for the subsequent evaluation. In first comparisons and demonstrations (binary image generation) Rowe et al. state, that the corresponding MSI data are able to produce recognizable fingerprint features for most of the fingers. This even holds for cases where the TIR image quality is severely degraded due to dry skin and/or other effects.

While the minutiae detection software can directly process a TIR image, the MSI data were at first preprocessed as mentioned before and then processed separately. After a preliminary review, the authors have discarded the 400nm images from the datasets because the corresponding illumination was too weak to achieve good fingerprints. Finally, the match results of the remaining six multispectral frames were averaged to a final match score.

In a biometrical performance study Rowe et al. discovered a dramatic difference in the performance of the TIR and multispectral fingerprint data. While the multispectral data produced an equal error rate (EER) of ~0.7%, the TIR sensor had an ERR of ~20%. Further

⁴ Neurotechnologija, VeriFinger 4.2

investigations have shown that the high error rate of the TIR data was found to be caused by certain study participants who produced an inordinate number of low-quality images due to dry skin. Therefore, a reassessment of the TIR data was done where the low-quality images are removed. This reassessment has shown a significant improvement of the ERR, which decreased to 1.8%.

4.3. Biometrics Based on Multispectral Skin Texture

As seen before, the utilization of multispectral data can improve the quality of biometrical fingerprint data. In [15], Rowe compared the performance of TIR and multispectral image based sensors with varying size. This comparison again is based on the above described sensor concept. In this case, also non-polarized illumination is considered. Here, the polarized and non-polarized multispectral illumination covers four wavelength ranges (430, 530, 630nm and white light). Using this sensor configuration, nine images are acquired, eight multispectral frames (polarized and non-polarized) and one TIR image.

The authors have compared their method with state-of-the-art work and exposed, that apart from differences in basis functions the prior work is always based on conventional fingerprint images with the already mentioned performance limitations. In contrast, they emphasize the main benefit of their work, which is the consideration of multiple images with information of the surface and also subsurface characteristics.

The method of dual-tree complex wavelet transform (DTCWT, cf. [1]), which produces complex coefficients, is applied to fuse the set of raw images into a single image. Further the DTCWT process was used to provide spectral-textual features of the multispectral data. Therefore, the coefficients from the third level of the DTCWT decomposition of the multispectral image stack were used as features. While the TIR image is included in the composite image, the author has omitted this image from the multispectral texture analysis, because of its high variability (e.g. skin moisture, contact based occurrences,

etc.). For the multispectral texture analysis, Rowe introduces a conjugate product P of coefficients at the decomposition level k .

$$P_{i,j}(x, y, k) = C_i(x, y, k)C_j^*(x, y, k) \quad (11)$$

where $C_i(x, y, k)$ is the complex coefficient for image i of the multispectral image stack and $C_j^*(x, y, k)$ is the conjugate of the corresponding complex value for image j [15]. The conjugate product P is referred as inter-image product and follows the inter-coefficient product of Anderson et al. [1], which has been shown to represent fundamental features while being translation invariant. For the purpose of classification, a feature vector is computed which consists of:

- real and imaginary components of all inter-image products for each unique image pair,
- isotropic magnitudes of coefficients (sum of absolute magnitudes over the directional coefficients),
- mean of directional coefficients values.

For the texture matching procedure, first a classification model is computed by using some calibration data as input for a Fisher linear discriminant analysis (FLDA) algorithm⁵. The differences in the feature vectors of test samples and classification samples are then projected onto the FLDA factors. The root mean square (RMS) values of the consequent projection results are finally accumulated as matching criteria. Small accumulation values mean high correspondence and therefore are treated as a match of fingerprints.

Again, a multi-person study, with 21 participants, was used to collect multispectral data. Each participant had to participate two times and in each visit data of four fingers were acquired three times. This procedure led in total to 504 multispectral datasets of 84 unique fingers. The data of the first visit were exclusively used as calibration data for the above described classification model. Consequently the data of the second visit were used as test samples.

⁵ Please note, that calibration- and also test data are normalized beforehand, so that they have a standard deviation of 1.

As mentioned before, the goal of Rowe was to investigate the feasibility of multispectral image based small-area sensors. For this reason, the analysis regions were varied in size (from 64x48 to 320x240 pixels). In the evaluation, the performance of the multispectral texture matcher and commercial minutiae-based fingerprint matching software⁶ were determined, whereas the software was utilized as control. Furthermore, twenty trials were realized. In each trial for each participant one fingerprint was randomly selected from the data of the first visit, which is used as calibration data. Then, data corresponding to six randomly selected fingers of the calibration data were selected as test samples from the data of the second visit.

The aggregation of the results has shown that the performance of both matchers is quite comparable for larger analysis regions. In the case of smaller regions the results indicate that the performance degrades significantly for the minutiae-based matcher, while the performance of the multispectral texture matcher is fairly high. As possible reason for the performance decrease of the minutiae-based matcher, the author supposes the following:

“This difference in performance is possibly due to the property of local consistency of the multispectral texture: skin proximal to the point of enrollment has approximately the same properties as the enrollment site itself. Therefore, placement-to-placement variation (which becomes more severe as the sensor size decreases) affects the texture matcher far less than the minutiae matcher.”

4.4. Summary

In this section, two approaches were presented to represent the feasibility of multispectral imaging techniques to improve the performance of fingerprint sensors. Among others, it was shown that conventional fingerprint sensors can be extended by multispectral image techniques to improve the performance in conditions where the performance of conventional TIR-based sensors is generally degraded by certain common occurrences. The assumed robustness is furthermore presented in the second work, where the results have

⁶ NEC, NECSAM FE4, ver. 1.0.2.0, PPC2003

shown that the performance is fairly high, even when the analysis regions decrease.

By utilizing multispectral data, the larger amount of information is considered as a major reason for the performance increase. While traditional fingerprint sensors capture single images of the skin surface, the pronounced multispectral sensor measures surface and subsurface characteristics of the skin, in different depth. This wealth of information leads to new research possibilities for novel or adapted methods (e.g. matcher) as exemplary shown.

5. Iris Analysis

In this section, we discuss the possibility of using iris as a human identification feature in multispectral data. The most iris recognition systems use exclusively the near-infrared (IR) range of the electromagnetic spectrum. In contrast to this, Boyce et al. [3] describe a preliminary study of multispectral iris analysis and take into consideration the iris information represented in both, the visible and IR portions of the spectrum. More specifically, they study the role of information represented in the IR, red (R), green (G), and blue (B) channel. To the best of our knowledge, by the time being there is no other work dealing with the problem of multispectral response of the iris tissue from a biometric perspective. Therefore, the content of this section is strongly based on the article by Christopher Boyce, Arun Ross, Matthew Monaco, Lawrence Hornak and Xin Li from the West Virginia University with the title *Multispectral Iris Analysis: A Preliminary Study* [3].

Boyce et al. [3] developed a system for multispectral iris analysis that takes classical steps of the pattern recognition processing chain, namely acquisition, preprocessing, feature extraction, classification, and evaluation. In the following, we will shortly describe their method and provide some general conclusions about the application of multispectral data in the area of iris analysis.

5.1. Acquisition and Preprocessing

Boyce et al. [3] used Redlake's MS3100 multispectral camera to acquire multispectral iris images. Their setup was equipped with two

separate ICX205AL sensors for the IR and R channels as well as one RGBICX205 sensor for the G and B channels. The resolutions in the different channels were not the same. While the IR and the R sensors provided images of size 1300x1040 pixels, the resolution in the G and B channels was thrice as low. For experiments, the authors acquired data from 24 individuals having different eye colors (5 samples per individual). Figure 3 show the intensity of iridal reflection across the four channels for different eye colors.

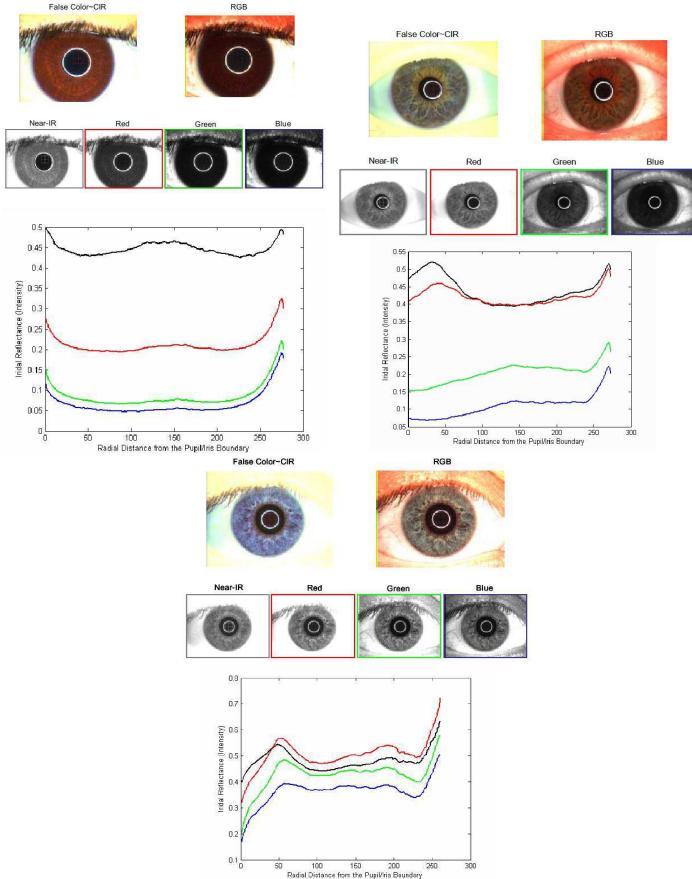


Fig. 3. Left: Example of a dark brown iris. The iris exhibits high iridal reflectance in the IR channel. The reflectance decreases significantly with wavelength; Middle: Example of a lightbrown iris. The iris exhibits high iridal reflectance in the IR and R channels. Reflectance decreases significantly for other wavelengths; Right: Example of a blue iris. The iridal reflection is comparable across all four channels;(Figure overtaken from [3])

In the preprocessing step, the authors of [3] perform automatic localization of the spatial extent of the iris structure by detecting its boundary in the image. First, the boundary between the pupil and the the iris was detected. Second, the boundary between the iris and the sclera was determined. Here, rather simple and known edge detection techniques were applied, since the focus of [3] is on iris image analysis. In Figure 4, results of the segmentation step in the IR, R, G, and B channels of a brown iris are depicted.

5.2. Feature Extraction and Classification

For feature extraction, the segmented iris was converted into the polar domain [3] and processed as a rectangular entity. Three Gabor filters with the same orientation and frequency but representing different scales were applied to observe the phase response of the filtered image. Each Gabor filter processed different region in the rectangular image since the level of iris detail typically degrades as one moves away from the pupil [3]. The resulting feature set consisted of an array of binary values and, therefore, the hamming distance was used as a metric in this feature space.

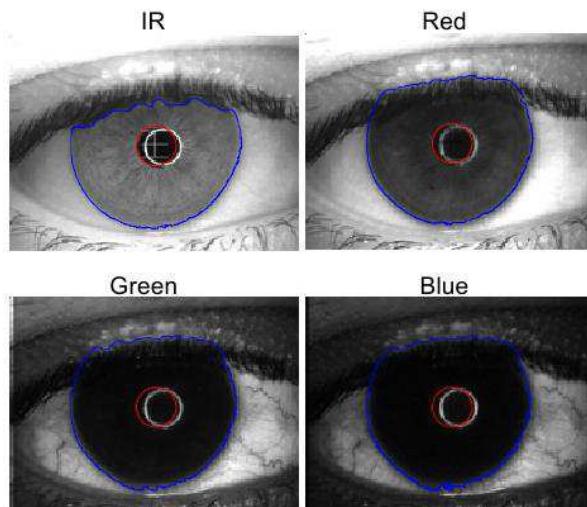


Fig. 4. Segmentation results in the IR, R, G, and B channels for a brown iris images (Figure overtaken from [3])

One of the objectives of the preliminary study presented in [3] was to discover, if there are any structural components of the iris (e. g., crypts, stroma, furrows, moles, freckles, etc.) that can be used for clustering of irides of different colors. To reduce the computational complexity, the image pixels were first averaged with a 4x4 smoothing filter and then down-sampled within individual spectral channels. Each pixel in the new image was then considered as a three dimensional vector with its elements corresponding to the R, G, and B channels. The K-means clustering algorithm using the Euclidean distance metric was then used to categorize the pixels.

In the first clustering experiment Boyce et al. [3] observed that for blue-colored irides, the different components of the iris were separated into multiple clusters. Then, the clustering procedure was repeated for iris images that were intentionally over-segmented by 100 pixels radially into the sclera. Here, the clustering results indicated the successful segmentation of the sclera, eyelashes and iris into separate components.

5.3. Summary

The purpose of the paper by Boyce et al. [3] was to highlight the potential of using multispectral iris information in recognition systems. It has been clearly stated that most commercial systems use only the near-IR wavelengths for iris analysis and completely disregard a possibility of employing multispectral information in this application domain.

The initial system presented in [3] uses well-known and very simple pattern recognition techniques. In our opinion, there is a lot of scientific potential in using multispectral data for iris-based human identification. However, the problem should be approached by a global optimization of all steps being part of the pattern recognition processing chain. In this context, a comprehensive suitability evaluation of different feature extraction techniques as well as classification schemes must be performed. E. g., the wavelet transform might be a good candidate for feature extraction in this application domain. The classification problem could be approached by clustering techniques based on cost function optimization [22]. Moreover, it

might be interesting to explore, whether learning-based (supervised) pattern recognition algorithms would bring better classification results in this application domain.

6. Conclusions

Multispectral imaging is an interesting tool in connection with personal identification. For all above considered applications, improvements of the conventional approaches can be observed. However, the motivation for these improvements is quite different.

Face recognition approaches are enhanced to cancel out uncontrolled acquisition conditions. Infrared is used to compensate poor or unclear illumination. Also, images with facial expressions are better matched. If the properties of the illumination and the acquisition system are known in advance, multispectral narrow bands can be used. Then, it is possible to normalize the images. This makes comparisons easier. Attempts to select these spectral bands which carry the most significant information have also been done successfully.

For palmprint and fingerprint biometrics, the acquisition conditions are determined due to the used scanners. Here, multispectral imaging aims at the improvement of the recognition performance. Secondary goal are spoof detection and reducing the size of the scanners. While (at least for finger prints) it could be shown that MSI can better handle a decreased region of interest, none of the approaches examined describes a working spoof detection method.

In Section 5, an initial work on multispectral iris analysis by Boyce et al. [3] was discussed and concluded. Most commercial systems use only the near-IR wavelengths for iris analysis and completely disregard a possibility of employing multispectral information in this application domain. Although the initial system presented in [3] uses well-known and very simple pattern recognition techniques, it shows clearly the benefits of using multispectral data for iris-based human identification. In our opinion however, the problem should be approached in a more systematic way. A global optimization of all steps being part of the pattern recognition processing chain is needed here. In this context, a comprehensive suitability evaluation of different feature extraction techniques as well as classification schemes

must be performed. E. g., the wavelet transform might be a good candidate for feature extraction in this application domain. The classification problem could be approached by clustering techniques based on cost function optimization [22]. Moreover, it might be interesting to explore, whether learning-based (supervised) pattern recognition algorithms would bring better classification results in this application domain.

All approaches have in common that once the information of the several spectral bands are fused they follow strictly the pattern recognition pipeline. As it seems, the first principle is to transform the multispectral information into an image which can be processed further by traditional, well-known methods. None of the examined methods considers the multispectral data as a whole. Thus, either the (traditionally used) feature extraction takes place at a fused image or features are extracted for each multispectral band separately.

Multispectral imaging has shown that it can outperform conventional approaches in terms of recognition rate. Therefore, it further remains an interesting research topic. However, multispectral methods developed in the future should focus more on the exploration of the whole spectral information. Also, for feature extraction, considering the spectral bands as whole could reveal new kinds of recognition algorithms.

Acknowledgments

Activities leading to this chapter have been funded by the German Research Foundation (DFG) in the context of the Research Training Group 1564 “Imaging New Modalities”.

References

- [1] Anderson R., Kingsbury N., Fauqueur J.: Robust rotation-invariant object recognition using edge-profile clusters, European Conference on Computer Vision, 2006
- [2] Bebis G., Gyaourova A., Singh S., Pavlidis I.: Face recognition by fusing thermal infrared and visible imagery, Image and Vision Computing, Volume 24, Issue 7, 1 July 2006, Pages 727-742

- [3] Boyce C., Ross A., Monaco M., Hornak L., Li X.: "Multispectral Iris Analysis: A Preliminary Study, Computer Vision and Pattern Recognition Workshop, p. 51, 2006
- [4] Chang H., Harishwaran H., Yi M., Koschan A., Abidi B., Abidi M.: An Indoor and Outdoor, Multimodal, Multispectral and Multi-Illuminant Database for Face Recognition Computer Vision and Pattern Recognition Workshop, p. 54, 2006
- [5] Chang H., Koschan A., Abidi B., Abidi M.: Physics-based Fusion of Multispectral Data for Improved Face Recognition, International Conference on Pattern Recognition, pp. 1083-1086, 2006
- [6] Chang H., Yao Y., Koschan A., Abidi B., Abidi M.: Improving Face Recognition via Narrowband Spectral Range Selection Using Jeffrey Divergence, Information Forensics and Security, vol.4, no.1, pp.111-122, 2009
- [7] Chui C.: An Introduction to Wavelets, Academic Press, London, 1992
- [8] Goldberg D.: Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading, MA, 1989
- [9] Guo Z., Zhang L., Zhang D.: Feature Band Selection for Multispectral Palmprint Recognition, Pattern Recognition, pp. 1136-1139, 2010
- [10] Gyaourova A., Bebis G., Pavlidis I.: Fusion of Infrared and Visible Images for Face Recognition, Computer Vision - ECCV 2004, Lecture Notes in Computer Science, T. Pajdla, J. Matas Ed., Springer Berlin / Heidelberg, 2004, Volume 3024/2004, pp. 456-468
- [11] Han C.-C., Cheng H.-L., Lin C.-L., Fan K.-C.: Personal authentication using palm-print features, Pattern Recognition, vol. 36:2, pp. 371-381, February 2003
- [12] Hao Y., Sun Z., and Tan T.: “Comparative studies on multispectral palm image fusion for biometrics, Lecture Notes in Computer Science – ACCV, 2007
- [13] Hao Y., Sun Z., Tan T., Ren C.: Multispectral palm image fusion for accurate contact-free palmprint recognition, Image Processing, pp. 281-284, 12-15 Oct. 2008
- [14] Rowe R.K., Nixon K.A., Corcoran S.P.: “Multispectral fingerprint biometrics, Annual IEEE Systems, Man and Cybernetics Information Assurance Workshop, pp. 14- 20, 15-17 June 2005

- [15] Rowe R.K.: "Biometrics based on multispectral skin texture, ICB 2007 - Lecture Notes in Computer Science, S.-W. Lee, S.Z. Li Ed., vol. 4642. Springer, Heidelberg, pp. 1144–1153, 2007
- [16] Rowe R.K., Uludag U., Demirkus M., Parthasaradhi S., Jain A.K.: A Multispectral Whole-Hand Biometric Authentication System, Biometrics Symposium, pp.1-6, 11-13 Sept. 2007
- [17] Rubner Y., Tomasi C., Guibas L. J.: The earth mover's distance as a metric for image retrieval. Int. J. Comp. Vis., vol. 40, no. 2, pp.99-121
- [18] Sirovich L. and Kirby M.: Low-dimensional procedure for the characterization of human faces, J. Opt. Soc. Am., pp. 519-524, 1987
- [19] Singh S., Gyaourova A., Bebis G. and Pavlidis I.: Infrared and Visible Image Fusion for Face Recognition, SPIE Defense and Security Symposium (Biometric Technology for Human Identification), Orlando, 12-16 April, 2004.
- [20] Singh R., Vatsa M., Noore A.: Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition, Pattern Recognition, vol 41:3, Part Special issue: Feature Generation and Machine Learning for Robust Multimodal Biometrics, pp. 880-893, March 2008
- [21] Singh R., Vatsa M., Noore A.: Hierarchical fusion of multi-spectral face images for improved recognition performance, Information Fusion, vol 9:2, pp. 200-210, April 2008
- [22] Theodoridis S., Koutroumbas K.: Pattern Recognition, Fourth Edition, Academic Press, 2008.
- [23] Turk M. and Pentland A.: Eigenfaces for Recognition, Journal of Cognitive Neuroscience, 3:1, pp. 71-86, 1991
- [24] Waymann J., Jain A., Maltoni D., Maio D.: Biometric Systems: Techn., Design and Performance Evaluation. Springer, London, 2005
- [25] Zhang D., Zhou Z.-H.: (2D)2PCA: Two-directional two-dimensional PCA for efficient face representation and recognition, Neurocomputing, volume 69:1-3, pp. 224-231, December 2005
- [26] Zhang D., Zhenhua G., Guangming L., Zhang L., Zuo W.: An Online System of Multispectral Palmprint Verification, Instrumentation and Measurement, IEEE Transactions on , vol.59, no.2, pp.480-490, Feb. 2010