We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Practical Imaging in Dermatology

Ville Voipio, Heikki Huttunen and Heikki Forsvik
Department of Signal Processing, Tampere University of Technology
Finland

1. Introduction

Visual observation has an important role in dermatology as the skin is the most visible organ. This makes dermatology a good candidate for utilizing digital imaging and automatic diagnostic tools. This chapter illustrates the use of 2D imaging in dermatology by presenting a method for automatization of common allergy testing. The so called "prick test" is the most common diagnostic test for different allergies. It involves measuring the skin response to percutaneously introduced allergens.

If there is a skin reaction against the allergen the blood flow is increased in the area and a wheal is formed. In physical terms, the color of the skin is changed, and a vertical displacement emerges. In this chapter we will concentrate on the color change (erythema) as we are only using 2D without the vertical displacement information.

Several approaches have been proposed to the skin erythema detection. The emphasis is often on detecting signs of melanoma and there exists a lot of literature on melanoma segmentation (Celebi et al., 2009; Gomez et al., 2008; 2007), but only a few studies of measurement of allergic reactions from 2D pictures (Nischik & Forster, 1997; Roullot et al., 2005). There are also some studies which have been performed with 3D imaging (Santos et al., 2008) or other specialized imaging hardware (Wöhrl et al., 2006).

We utilize the inexpensive commonplace 2D digital photography. The key challenges in this approach are the transformation of the color image into a maximal-contrast single-variable (grayscale) image and the interpretation of the wheal dimensions from this transformed image. This chapter is organized so that section 2 discusses the effect of the imaging hardware and camera settings on the medical and scientific imaging. After that we present our algorithm for segmenting the wheal area in Section 3, which is based on our earlier paper (Huttunen et al., 2011). Finally, Section 4 discusses the results.

2. Imaging

The easy availability and high quality of digital cameras made them attractive for medical and other scientific use. Purpose-built scientific digital cameras carry typically a tenfold price tag, and the pace of development is very fast in the mainstream digital cameras.

While the actual performance of the inexpensive cameras is usually very good, their image processing chain is not well documented. Scientific cameras are supplied with photometric specifications, and they do not perform any unspecified processing to the image.

Photography-oriented equipment very seldom gives any photometric specifications, and the cameras may perform a lot of unspecified image processing.

The low cost of ordinary digital cameras opens up a large number of new opportunities in medicine. In order to avoid some common pitfalls associated with their use, one should be aware of the typical image processing in the camera. The actual processing varies from one camera to another, and the image processing methods are not published by the manufacturers, but the following sections outline the fundamental limitations and give an overview of the signal processing in a typical digital camera.

The discussion below will concentrate on digital cameras with interchangeable optics (typically SLR, Single Lens Reflex cameras) if not indicated otherwise. The use of compact cameras is usually not recommended in medical or scientific work; SLR cameras provide much better imaging properties and are still affordable.

2.1 Signal path

Finding a simple and accurate definition for a digital camera is difficult, as the use of a camera may range from art to measurement. In the following discussion the digital camera is viewed as *a measurement device measuring the spectral intensity distribution of light arriving from different spatial angles towards the camera*. Thus, in an ideal camera the spectrum of light arriving from each angle to the film plane is recorded.

Evidently, the definition above gives an infinite amount of information and is in contradiction with both the real world situation and physics. In order to understand the practical limitations, we may have a look at the signal path in a digital camera and the limitations imposed by each step in the path (see figure 1).

INCOMING LIGHT
- shot noise

COLOR FILTER
- loss of intensity
- loss of resolution

READ-OUT AMPLIFIER
- read-out noise

OBJECTIVE LENS
- loss of resolution
- loss of contrast
- reflection losses

DETECTOR ARRAY
- thermal noise
- quantum efficiency <1
- defective pixels
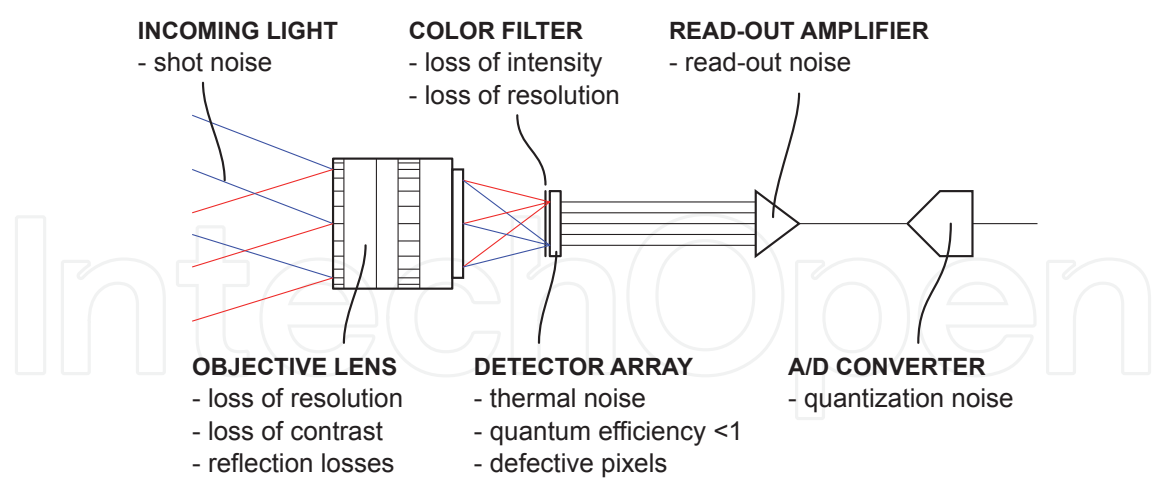
A/D CONVERTER
- quantization noise

Fig. 1. Signal path in a digital camera and sources of loss of information associated with each step.

The first and a very fundamental reason for loss of information lies in the quantum nature of light. As the light collected by the optical system consists of a number of photons, there is always some statistical noise ("shot noise"). Shot noise is usually the dominant noise source in

photography. The signal-to-noise ratio limited by the shot noise is inversely proportional to the square root of the intensity of the light falling onto the detector.

2.1.1 Objective lens

The purpose of the optical system is to direct photons arriving from a certain direction onto the detector plane. There are two physical limitations related to the optical system; limited depth-of-field and diffraction limit. If the optical system is to collect more light, its light collecting area (usually a circular aperture) has to be larger. Inevitably, large aperture makes a shorter depth-of-view, i.e., only objects at a certain limited distance range are in focus (see section 2.3.2). On the other hand, if the aperture is small, the image becomes diffraction limited due to the wave nature of light. In practice, even moderately low cost digital camera systems are often diffraction limited in their behaviour.

The diffraction limit for an optical system is given by the Rayleigh's criterion:

$$d = 1.22\frac{\lambda f}{D}, \tag{1}$$

where d is the smallest resolvable detail on the image plane, λ wavelength, D the size of the entrance pupil (aperture), and f the focal length of the objective. For a typical digital camera with a typical objective f/D is in the range of 2..20. The wavelength of green light is roughly 500 nm, and the pixel size of a SLR digital camera is typically approximately 5 micrometers, and hence the resolution becomes easily diffraction limited when $f/D > 8$.

The discussion above is somewhat simplified in two respects. First, the Rayleigh's criterion is somewhat arbitrary, it actually refers to the distance between the center and the first minimum of the Airy disk (a diffraction pattern arising from an aperture), and has its roots in astronomy. Two sharp spots of light can be resolved even when they are closer than the Rayleigh's criterion, so the criterion is more a rule of thumb than an exact measure (Hecht, 1987). Second, in practice the objective may not be diffraction limited at or close to its largest aperture setting.

The optical system may also lose some light due to absorption and reflection losses. While the absolute amount of light lost this way is usually rather small, these losses may be seen as uneven illumination across the field (usually less light in the corners) or loss of contrast when there are very brightly illuminated areas in the image.

2.1.2 Color filters

Virtually all consumer digital still cameras are color cameras. As the light-sensitive pixels in the image detector cannot detect the color of the incoming light, a color filter is required[1]. The color filter represents a major loss of light, as the photons of mismatching energies ("wrong color") are absorbed.

The color filters are laid out so that neighboring pixels have different filters in front of them. There are numerous different filter patterns, both the layout and the actual colors of the filters

---

[1] There is one notable exception to this rule; Sigma Corporation Foveon X3 sensors have three photosensitive sensors overlaid on top of each other, which in principle eliminates the filter and losses associated to it. Also, many digital video cameras employ a color-separation beamsplitter which directs different wavelengths to three separate image sensors.

vary. Probably the most common of these patterns is the Bayer pattern where a 2x2-pixel block has one red, one blue, and one green filter (see Figure 2). The choice of the filter pattern and color is a compromise between spatial resolution, light collecting efficiency, and color reproduction. There are even some cameras which have four color filters (RGBE, E for emerald) to increase the color sensitivity.

It should be noted that whatever the filter pattern, the color reconstruction (demosaic) filter is not a trivial one. If the maximum resolution is desired, the filter needs to identify edges by combining information from several pixels. This requires non-linear processing and is prone to creating artefacts.

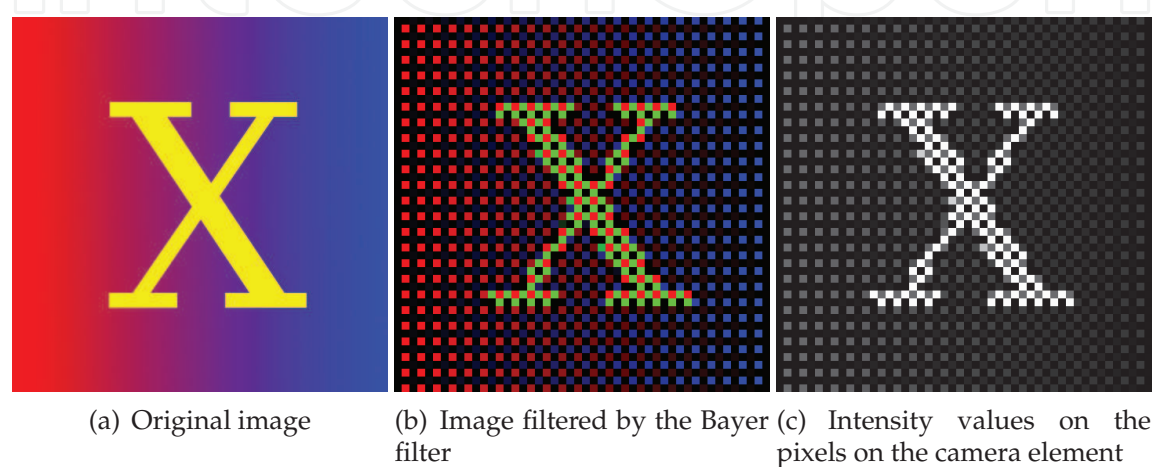| (a) Original image | (b) Image filtered by the Bayer filter | (c) Intensity values on the pixels on the camera element |

Fig. 2. Color reproduction in a Bayer filter.

It is important to note that color cameras are essentially very crude three-channel filter spectrometers; the full spectrum of visible light is reduced to three values. The spectral properties of the color filters are usually not disclosed, and in this way the color information given by a digital camera can usually be interpreted only in a relative manner.

Color is not a physical property of light, it is a property of the human visual system. There are three different types of color sensitive cone cells in the human retina. One cell type is sensitive to longer wavelength (red), the other to medium wavelengths (green), and the third type to short wavelengths (blue). The RGB system for color pictures is adopted with the idea that each of the primary colors corresponds to one type of cone cells.[2]

If the transmission properties of the color filters in a digital camera differ from those of the retinal cone cells, the color space as seen by the digital camera is not the same as seen by the human eye. There may be metameric colors (colors which look the same to a human observer but have different spectral distribution) which a camera may be able to distinguish, and vice versa. There is no simple way to describe these relations without knowing the actual filter characteristics of the color filters in the camera, and even then the practical implications are not necessarily evident.

---

[2] While a RGB image sensor can cover the full range of colors seen by the human eye, the same does not apply to reproducing the color. Mixing light of any three colors can never produce the full range (gamut) of colors which the human eye can see.

2.1.3 Image detector

After going through the color filters, the light is detected by a CCD (Charge-Coupled Device) or a CMOS (Complementary Metal Oxide Semiconductor) array camera. The difference between the two technologies is rather subtle, in both devices a photon hitting a detector cell has some probability of generating a charge carrier which is trapped in the pixel. This probability depends on the sensor and the wavelength of the incident photon and is called the quantum efficiency (QE).

The actual quantum efficiencies of real-world cameras are trade secrets, but Farrell et al. (2006) suggests that the quantum efficiency of a digital SLR pixels is around 0.65, but as the fill factor (proportion of active pixel area to the total sensor area) is only around 50 %, the total QE is approximately 1/3.

The pixel size limits the spatial resolution of the image. However, this is not usually the most limiting factor, as the physical and practical limits of the optics already limit the resolution. Also, making small pixels may reduce the fill factor and increase the relative read-out noise. The choice of pixel size is a compromise between dynamic performance and resolution of the sensor (Farrell et al., 2006). In practice, there are very few situations in medical imaging where a larger number of pixels would be beneficial.

Once the light has been captured by a sensor pixel, it is read out, amplified and quantized by an AD (analog-to-digital) converter. The pixel itself has some temperature-dependent thermally induced noise. Usually this noise is significant only with long exposure times (such as in astrophotography). The pixel readout noise is a source for constant noise which is not dependent on the illumination. Thus, it is usually significant only with low illumination levels (deep dark areas or very low light).

Also, the conversion gain from stored charge carriers to a voltage signal varies from one pixel to another. There may even be hot (always at full output) or dead (always at zero) pixels due to manufacturing defects in the sensor array. Hot pixels may be such that they show only at very long exposure times.

2.1.4 AD conversion

While the number of bits in the AD converter seems to be a an important marketing feature for some digital cameras, the quantization noise of the AD converter is typically far below the other noise sources (readout in the dark area, shot noise in the highlight areas). The maximum dynamic intensity range offered by a 14-bit AD converter is $1 : 2^{14} = 1 : 16384$, whereas the typical maximum dynamic range of the sensor limited by other factors is roughly 1:2000.

Practically all digital cameras offer the possibility to change the sensitivity (ISO number) of the camera. However, this choice does not change the sensitivity of the photosensitive elements. Instead, it either changes the amplification of the analog signal between the camera element and the AD converter or is purely a digital multiplication. As these operations do not change the fundamental noise sources (shot noise, thermal noise, readout noise), it is better to increase the illumination instead of increasing the sensitivity setting.

In this signal chain from the spatial distribution of light to bits from the sensor it is important to note that many of the limitations are fundamental physical limitations (shot

noise, depth-of-view, diffraction limit). The camera technology is very good even at today's level, and the room for improvement is rather limited.

2.2 Digital camera signal processing

Before the image from the sensor is available as an image file, it undergoes several digital processing steps. There are at least three different goals in this processing:

1. Correcting image defects

2. Improving the visual appearance of the image

3. Making the image technically easier to use

No image processing can increase the amount of information in the image. Some of the image processing steps may be useful in scientific imaging, but others may make the image less useful in further image processing. At this point the purpose-built scientific cameras are usually more faithful to the actual scene. The image processing steps and algorithms in each individual camera type are proprietary, but there are some general steps which are almost unavoidable:

• Pixel-to-pixel non-uniformity (fixed noise) compensation

• Dead pixel elimination

• Demosaicing (color reproduction)

After these processing steps the result is a RGB image in a linear color space. This is usually the most useful image form for further image processing. Many digital cameras provide such an image under the name "RAW". However, the raw image format is manufacturer-specific, usually does not have any demosaicing applied, and may or may not compensate for dead pixels. The RAW formats are seldom officially published, but there are utilities such as `dcraw` (Coffin, 2011) which can be used to convert the images to some format easier to handle with image processing software.

One attempt to standardize the varying RAW file formats is the DNG ("digital negative") format by Adobe (Adobe Systems Incorporated, 2009). The DNG specification gives an overview of different aspects of RAW images. Despite this attempt, using the RAW images may be slightly challenging in practice.

In everyday photography, the images are usually processed in-camera with at least the following steps:

• Noise reduction

• Illumination correction (white balance)

• Sharpening

• Color space conversion to a non-linear one

• Compression

The purpose of this processing is to make the image look visually pleasing and occupy small space. Unfortunately, these steps reduce the information available in the image. The goal of

color space conversion and compression is to reduce the number of bits without removing visually important information. For example, the very commonly used JPEG (ECMA, 2009) image compression reduces the chrominance (color) information more than the luminance (brightness) information, as the human visual system is more sensitive to luminance changes.

The camera image processing may be aware of certain optical properties of the objective lens, e.g., chromatic aberration, geometric aberrations, vignetting (loss of light at the corners), and may perform corrective operations.

For machine vision applications, there are two possibilities when choosing the file format. If a single camera type is used, then using RAW files of that camera preserves as much information as possible. If there is no control over the camera type, then setting the camera to produce JPEG files with as little processing (sharpening, noise reduction, compression) as possible gives a reasonably standardized result.

It should be noted that even the RAW images from the camera may be compressed images. An analysis of `dcraw`'s source code reveals, for example, that some RAW files have the AD outputs further quantized. Whether or not such operations actually reduce any useful information from the image is unknown, as their effect may remain below the noise level.

The example images used in this chapter have been compressed in-camera with a moderate JPEG. This approach is not optimal, and some of the noise visible in Figure 6 is likely to be compression artefacts. Note that all color space transformations tend to emphasize the compression noise, because JPEG is designed for good visual appearance in the RGB color space.

One additional concern is storing the images for long-term use. JPEG images can most probably be read a long time in the future as it is the most common photographic image file format at the moment. The less common RAW formats may be difficult to interpret in the decades to come, as none of the formats has gained significantly larger popularity than the others. In medical applications the images may be stored in the DICOM (NEMA, 2011) format. This does not change the work flow in the digital camera in any way, and the DICOM format offers a number of different lossy and lossless data encoding schemes, whose suitability depend a great deal on the imaging application itself. However, the use of DICOM does not necessarily mean the images will be completely future-proof, as the standard and its use evolve.

2.3 Photographic setup

Designing a suitable photographic setup is very important in scientific or medical photography. The arrangement of the photographic setup has usually much more impact on the usability of the image than differences between different cameras.

The purpose of photographic imaging in dermatology is to measure the spectral reflectance of skin at different points in the area of interest. The accuracy of this process depends on a large number of parameters related to the illumination, setup geometry, optics, camera settings, and the camera itself.

A diagram of the photographic setup used in our project is shown in figure 3. The forearm of the patient is supported so that it does not move during the test. The support is built to
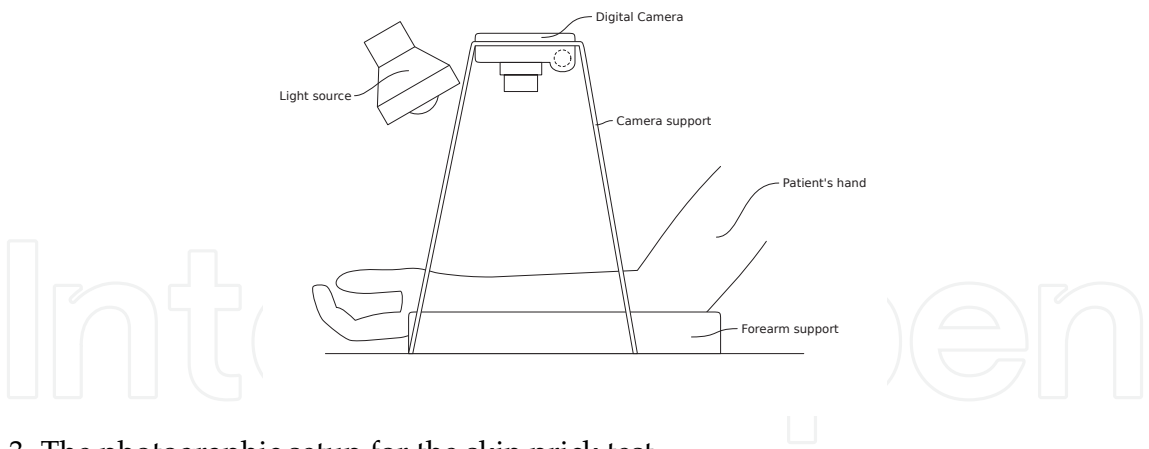
Fig. 3. The photographic setup for the skin prick test.

eliminate unnecessary muscular strain, as the test takes a half an hour to complete. The light falls from the direction of the camera to eliminate shadows. While the basic structure of the test setup is well-suited to the application, the illumination can be improved. Especially, the spectral properties of the light source were suboptimal, as there was little short wavelength (blue) radiation in the illumination.

2.3.1 Illumination

Illumination has an extremely important role. There are three main aspects to be taken into account in designing the lighting:

1. amount of light
2. spectral properties of the illumination
3. geometric properties of the illumination

In practice, the requirement for a correct amount of light translates to the requirement of having enough light. Insufficient illumination decreases the amount of information in the photograph either as a loss of resolution or as an increase of noise. (In some rather rare cases one should also pay attention to the effect of the light or thermal radiation onto the tissue illuminated.)

The way colors are formed in a digital camera actually means that the camera measures the product of the spectrum of the illumination, spectral reflection of the target, and spectral transmission of the camera color filters. This process naturally loses most of the spectral reflection information of the target, and accurate measurement of color-related quantities is difficult in digital photography.

Table 1 summarizes the typical spectral properties of some illuminants. The variation within each technology is usually significant, and if reliable color reproduction is required, the light source has to be chosen and calibrated carefully. It should also be noted that changing the camera type changes the camera filter absorption spectrum and thus the color reproduction.

In most cases the color of the illuminant can be sufficiently compensated for by the white balance compensation. The light reflected off a white or neutral gray target should result

Lamp type	Spectral properties
Incandescent lamp	Smooth spectrum, low intensity in the blue (short wavelength) region (tungsten halogen lamps have more blue emission)
Fluorescent tube	Usually a spectrum consisting of a large number of separate narrow peaks
Photographic flash	Several rather narrow peaks
Monochromatic LED	A single peak with a typical peak width of 20 nm
White LED	Either a single narrow blue peak and a wide yellow peak ("white LED") or three narrow separate peaks (combination of R, G, and B LEDs). The color may depend on the radiation angle.

Table 1. Typical spectral properties of different light sources

in a RGB value with equal amount of each of the components. If this is not the case, the components have to be adjusted by a multiplication:

$$
\begin{pmatrix} R_{corr} \\ G_{corr} \\ B_{corr} \end{pmatrix} = k \begin{pmatrix} 1/R_{gray} & 0 & 0 \\ 0 & 1/G_{gray} & 0 \\ 0 & 0 & 1/B_{gray} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}, \tag{2}
$$

where X is the uncorrected value for color X, X_{gray} the uncorrected value for the gray surface illuminated with the same illuminant, X_{corr} the corrected value, and k a normalizing factor. If the automatic white balance setting offered by the camera is used, the camera algorithms try to find the optimal correction coefficients. This process is based on typical photographic scenes, and it may result in unexpected correction results in scientific photography. It should be noted that the white balance correction is often quite significant especially in the blue channel.

Most cameras try to convert the information from the image sensor to the commonly used sRGB color space. This conversion usually involves using a 3x3 matrix multiplication from the camera RGB to the sRGB color space. The conversion may be beneficial for human viewing, but after it has been performed, the white balance equations above do not hold true. For an in-depth discussion of white balance, see Viggiano (2004).

Film-based medical photography used to have some applications with monochrome film and an external filter to enhance the contrast of specific features. In general, this method cannot be substituted with digital color photography and digital post processing, as the transmission properties of the filters do not match those of the color filters of the camera. The solution is to use a filter in front of the color camera, filter the light source, or use a (quasi-)monochromatic light source.

The geometric properties of the illumination have a lot of impact on the contrast of the image. The illumination may create shadows or specular reflections (bright spots) which are usually undesired, but in some cases they may be useful in detecting feature outlines or surface normal directions. In general, a point-like light sources give more contrast and more shadows and specular reflections than area sources.

There is very little in the camera or in the image processing that can be done to compensate for a bad illumination geometry. However, if shininess or specular reflections are the only problems in otherwise good photographic setup, a polarizing filter may help, as it blocks the light reflected at certain angles.

2.3.2 Camera settings

There are essentially four independent settings in the camera which can be adjusted:

1. focal length
2. focus
3. aperture
4. exposure time

The focal length together with the image sensor size determine the field of view. In a given photographic setup the position of the camera is usually fixed, so that an objective with a suitable focal length needs to be used. In general, objectives with fixed focal length exhibit better optical performance than zoom objectives. The optimal distance between the target and the camera depends on the application, but usually very short focal lengths (wide field of view) are to be avoided, if possible. Figure 4 illustrates the effect of focal length on the perspective of the image.
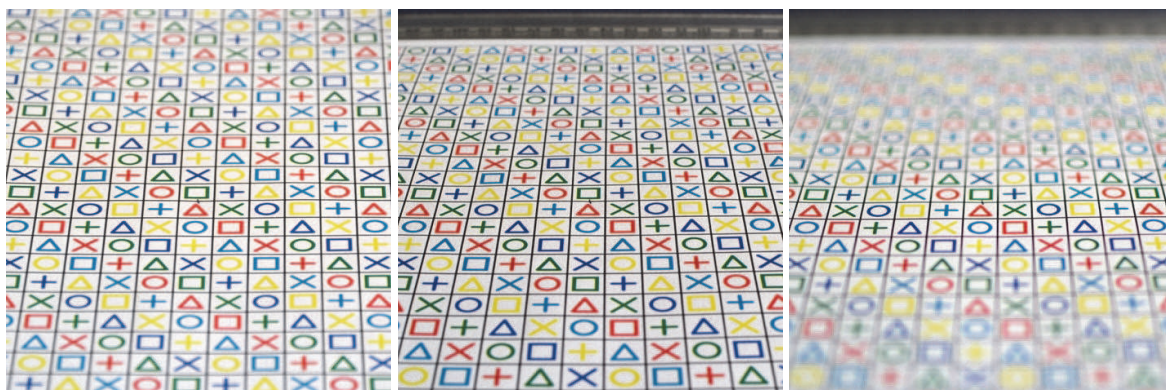
(a) Focal length: 105 mm, $f/7.1$ (b) Focal length: 35 mm, $f/7.1$ (c) Focal length: 35 mm, $f/1.8$

Fig. 4. Effect of focal length and aperture changes. The perspective distortion is much smaller at a longer focal length, and the depth of field is shorter with larger aperture (small f-number). The target and the imaging angle are the same in all images.

The digital cameras offer a selection between automatic or manual focus setting. If the imaging setup is fixed, the manual focus setting is preferred, as the autofocus algorithms are tuned to everyday photography, and the scenes in medical or scientific imaging are different.

The aperture (or entrance pupil) of the optics describes the size of the light collecting area of the objective. A large aperture collects more light and is thus useful in low-light situations. However, the depth of field (useful focus range) is very short at a large aperture (see figure 4). Also, at the large end of the aperture range the image resolution may suffer due to practical limitations. On the other hand, very small apertures have more diffraction, and this may

slightly deteriorate the image resolution, as well. The optimal choice of the aperture depends on the application, and finding it may require some experimenting.

It should be noted that comparing the actual light collecting aperture of different cameras is not straightforward. The aperture size is usually given as the f-number (or $f/\#$), so that the actual aperture size in physical units is the focal length divided by the f-number. Thus, a large f-number indicates a small aperture, and the f-number itself is not meaningful unless the focal length is known. Even in digital cameras with interchangeable lenses the aperture area may vary in the range of 1:4 with the same field of view and aperture number.

The exposure time is perhaps the simplest adjustment. The longer the exposure time, the more photons reach the imaging element. Usually the shortest exposure times available in digital cameras are below one millisecond, and the longest times in the order of dozens of seconds. In practice, the short end of the exposure range is not very useful in medical photography, as there very seldom is enough light. On the other hand, exposure times above 1/60 s are prone to movements in the image. The amount of movement depends very much on the actual application, sometimes much longer exposure times are useful.

In some cases the dynamic range of the scene surpasses that of the digital camera. The dynamic range of a digital camera is typically approximately 1:2000. If this is not sufficient for some reason, it is possible to take several shots of the same scene with different exposure times. This technique is called bracketing, and the imaging method carries the acronym HDR (High Dynamic Range). The Achilles' heel of HDR is the time difference between different shots, which usually makes it impossible to apply to moving objects.

3. Wheal shape and size recognition

The skin prick test is widely used in allergy testing. As the test results in local skin reactions (wheals) around a well-defined test site, it lends itself well to computer-based interpretation.

3.1 Medical background

The skin prick test is a well-known and well-established method for quantitative measurement of allergic reactions (Oppenheimer et al., 2011). An allergen is introduced percutaneously; a drop of allergen solution is dropped onto the skin, and the skin is then punctured by a small blade designed for this purpose. If there is an allergic reaction, a wheal will emerge.

3.1.1 Skin reactions

If an allergen is introduced percutaneously, and there are IgE antibodies to that allergen, an inflammatory reaction will arise. Two different types of cells are involved in the process, basophils and mast cells. These cells release chemicals associated with inflammation, such as cytokines and histamines. These chemicals then mediate processes such as vasodilatation (expansion of the blood vessels) and increased permeability of the blood vessel walls.

The result of these processes is redness and heating due to increased blood flow and swelling due to leakage of fluid into the surrounding tissues. In the case of the prick test the reaction is localized around the prick site and produce an approximately circular wheal. In practice, the shape of the wheal may vary a lot depending on the structure of the surrounding tissue.
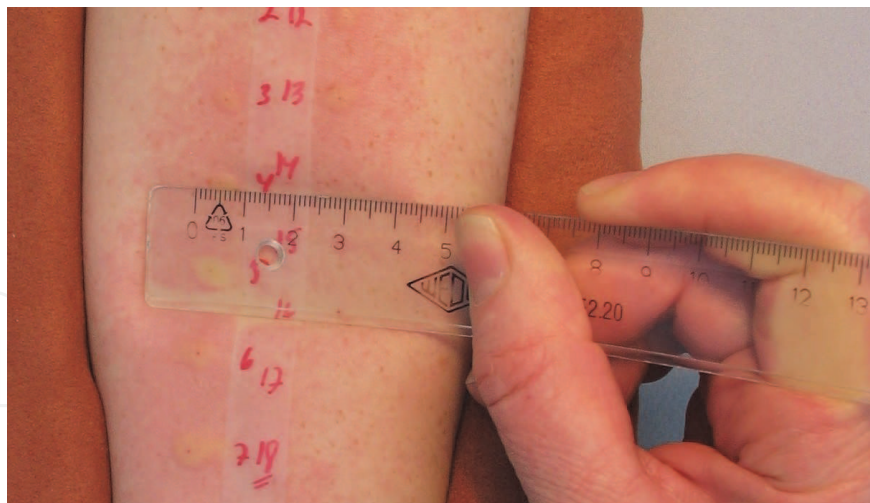
Fig. 5. Manual allergy measurement.

The prick test has some shortcomings. The magnitude of the skin response varies from one patient to another. For example, with young patients the size of the wheal becomes larger as the patient grows (Meinert et al., 1994). Also, the correlation between the actual allergic reactions and the skin reactions is not always very good, and there are several practical factors which may change the results (Haahtela et al., 2010).

3.1.2 Prick testing in practice

In practice, the prick test is used to test several allergens at the same time. To facilitate this, a suitable even skin area is required, the most common such area being the inner forearm. The number of allergens tested at the same time varies, but typically the inner forearm can carry 20 pricks.

The results are read some 15 to 20 minutes after the prick. The time is sufficient for the reaction to emerge but not long enough to let the symptoms fade. There is some evidence that if the time is too long, the reliability of the test may suffer (Seibert et al., 2011). However, the actual development of the inflammation reaction as a function of time is not generally known, as there have not been any suitable tools for measuring it. One of the aims of our research is to introduce these tools, as a series of photographs reveals the development of the wheal as a function of time.

There are different readout methods in use worldwide. In many cases the results are evaluated pseudo quantitatively by visual inspection only. Seibert et al. (2011) argue that the most reliable way of testing is to measure the size of the wheal, as the visual estimation has been shown to be highly variable.

When the wheal size is measured, there are several different ways of doing it. Traditionally, the practitioner uses a ruler to manually measure the size of the wheal. The test procedure assumes an elliptic shape, with possible elongated branches (called *pseudopodia*) disregarded and the result of the measurement is the mean of the major and minor axes of the imaginary ellipse (Santos et al., 2008). There is also a good correlation between the length of the long axis and the area of the wheal, but the mean method should yield even better results.

The practice of measuring the wheal vary. The practitioner may use a transparent ruler and slightly press the wheal to see the outline, or she may try to use the color change only. An illustration of the measurement is in Figure 5, where the application of pressure can be seen. Our computerized interpretation method uses the color change only, and the wheal is always interpreted as an ellipse.

One of the challenges in skin prick testing is the difference of reactions between different individuals. A common way to account for this is to use histamine as one of the test allergens. It is not an allergen, but it always causes an inflammatory reaction, which can be used as a reference of the individual reaction. Saline solution can also be introduced as a zero control which should not cause any reaction. We used both a zero reference and histamine reference in our study.

3.2 Wheal extraction by grayscale transformation

The problem of finding the area within a wheal in a photograph carries some challenges. The color of the skin varies from one person to another, and the skin does not have an even color. The wheal edges are not sharp, so a threshold has to be defined. Naturally, the first step in finding out the wheal area is to convert the RGB image into a single-variable (grayscale) image. Preferably, the method should be adaptive, as the range of skin colors and illumination variations is large.

Among the earlier studies in the field, Roullot et al. (2005) considers seven well known color spaces and compares the separability of the reaction from the background using a training database. As a result, they discover that the optimal dimension among the color spaces is the a^*-component of the $L^*a^*b^*$ color space. Using the extracted a^*-component, they use simple thresholding for segmenting the wheal.

Nischik *et al.* (Nischik & Forster, 1997) also discover the $L^*a^*b^*$ color space most suitable for the wheal segmentation and use the standard deviations of the L^* and a^* components as the features for classification. The classifier is trained to separate between foreground (the wheal) and the background (healthy skin) using manually generated training data. The classifier output determines directly the boundary between the two regions.

Recent work by Celebi *et al.* consider finding optimal color transformation for extracting the foreground (Celebi et al., 2009). Although the paper concentrates on melanoma segmentation, the principle is applicable for other skin diseases, as well. The paper searches for optimal linear combination of the RGB-components, such that the output maximizes the separability of the foreground and background. The foreground and background are determined in each iteration using Otsu thresholding. Thus, the algorithm iterates all projections defined on a finite grid, and tests their performance by measuring the Fisher ratio of the foreground and background (which are determined using Otsu thresholding).

The method of Celebi et al. (2009) is an unsupervised method, which attempts to find the best projection without any user manual assistance. However, in our work we study the case, in which the user points the approximate center of the wheal. From the practical point of view this is acceptable, because it requires less work than the manual measurement. However, we plan to automate the detection of the wheal location in the future. Note also that in the
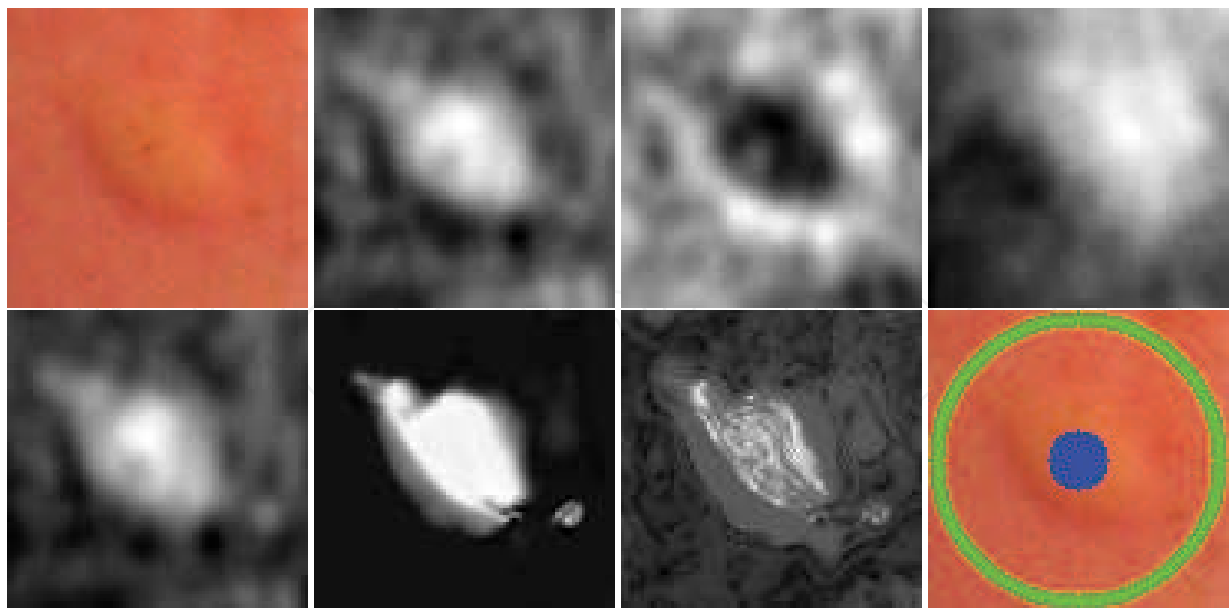
Fig. 6. The projection of the wheal in RGB color space. The original RGB image is shown in Figure (a), and (b) shows the hue component, (c) the a^* component of the $L^*a^*b^*$ color space and (d) is the projection proposed in Celebi et al. (2009). Optimal Fisher discriminant projection is shown in Figure (e), and the results of its kernelized version are shown in figures (f) and (g) using RBF kernel with bandwidth σ selected using Silverman rule of thumb and with fixed $\sigma = 0.5$. The training sets are obtained from areas shown in (h), where the blue center is the foreground sample region and the green circle is the background sample region.

temporal direction, clicking the last image in the time series is enough for determining the wheal location in all pictures, if temporal motion compensation is used.

The key problem when searching the borders of the wheal is the poor contrast between the wheal and skin. An example of a wheal is illustrated in Figure 6 (a). Although the wheal borders are barely visible, the shape becomes highlighted when mapped into grayscale in a suitable manner. Well known mappings for skin color processing include the hue component of the HSV color space (Figure 6 (b)) and the a^* component of the $L^*a^*b^*$ color space (Figure 6 (c)). In all projections, we have smoothed the RGB image by convolution with a disc shaped window of radius 5. However, these are more or less arbitrary, and variability in skin color and allergic reaction strength may decrease their applicability. Instead, training based projections may improve the separation further, and make it more invariant for all patients. An unsupervised method for finding a well-separating projection in terms of the Fisher criterion was proposed by Celebi *et al.* (Celebi et al., 2009), whose result is shown in Figure 6 (d). In this case the coefficients are 1, -0.1 and -0.3 for red, green and blue channels, respectively.

Optimality of the grayscale projection can be studied assuming that we know the approximate location of the wheal. This way we can construct training sets consisting of the wheal area and the surrounding healthy skin, denoted by $\mathcal{S}_1$ and $\mathcal{S}_0$, respectively. With the training sets we can seek for optimal separation in the RGB space in a supervised fashion.

The training set is acquired as follows. When the user has pointed the approximate location of the center of the wheal, a set of RGB values is obtained from the neighborhood. In our

experiments, the training set of the wheal ($\mathcal{S}_1$) is obtained inside the circular neighborhood with the radius of 10 pixels. The training set of the healthy skin ($\mathcal{S}_0$) is acquired from pixels that are far away from the center. We have used all pixels located at a radius between 45 and 50 pixels from the center, as illustrated in Figure 6 (h).

The natural tool for optimally projecting the three-dimensions to grayscale is the *Fisher Discriminant*, Fisher (1936). Fisher discriminant finds the projection dimension **w** that maximizes the separability of the classes in terms of the ratio of the between-class-variance and within-class-variance; i.e., the so called Fisher ratio:

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}, \tag{3}$$

where $\mathbf{S}_W \in \mathbf{R}^{3 \times 3}$ and $\mathbf{S}_B \in \mathbf{R}^{3 \times 3}$ are the within-class and between-class scatter matrices, respectively. It can be shown that the optimal direction **w** is given by

$$\mathbf{w} = \mathbf{S}_W^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0), \tag{4}$$

where $\boldsymbol{\mu}_1 \in \mathbf{R}^3$ and $\boldsymbol{\mu}_0 \in \mathbf{R}^3$ are the sample means of $\mathcal{S}_1$ and $\mathcal{S}_0$.

An example of the result of the Fisher discriminant projection is shown in Figure 6 (e).

The FD is a special case of the so called *Kernel Fisher Discriminant* (KFD) (Mika et al., 1999; Schölkopf & Smola, 2001), which is a kernelized version of the standard FD. As all kernel methods, the KFD implicitly maps the original data into a high-dimensional feature space and finds the optimally separating manifold there. Using the implicit mapping via the *kernel trick*, the explicit mapping can be avoided, which allows calculating the FD even in an infinite-dimensional space.

The kernel trick enables better separation between the two classes, which is important because the foreground and the background are typically not linearly separable in the original RGB space. For example, it might be that the foreground is in the middle of two background color regions in the RGB space. One consequence of this is the fact that most authors prefer the $L^*a^*b^*$ color space, because the classes are better linearly separable there. However, the kernel trick makes the color space transformation less significant due to the transformation into a higher dimensional space, where the classes will be better separable almost regardless of the original color space.

Denote the foreground samples by $Y_1 = \{\mathbf{x}_1^F, \mathbf{x}_2^F, \ldots, \mathbf{x}_{N_F}^F\}$, and the background samples by $Y_0 = \{\mathbf{x}_1^B, \mathbf{x}_2^B, \ldots, \mathbf{x}_{N_B}^B\}$, where each sample is a 3-dimensional vector in the RGB space: $\mathbf{x}_k^{F,B} = (r_k^{F,B}, g_k^{F,B}, b_k^{F,B})^T$.

3.3 The Fisher discriminant projection

After the samples from the two classes have been collected, the optimally separating linear transformation is given by projection onto the vector **w** defined by the Fisher discriminant as follows,

$$\mathbf{w} = (\mathbf{C}_1 + \mathbf{C}_0)^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0),$$

where $\mathbf{C}_1 \in \mathbf{R}^{3\times3}$ and $\mathbf{C}_0 \in \mathbf{R}^{3\times3}$ are the sample covariance matrices and $\boldsymbol{\mu}_1 \in \mathbf{R}^3$ and $\boldsymbol{\mu}_0 \in \mathbf{R}^3$ the sample means of the foreground and background samples, respectively.

3.4 The kernel Fisher discriminant projection

The Fisher discriminant projection is a special case of the kernel Fisher discriminant projection, occurring when the implicit mapping function is $\boldsymbol{\Phi}(\mathbf{x}) = \mathbf{x}$. The use of more complicated mapping functions allows more complicated separations for the classes. For arbitrary mapping the Kernel Fisher Discriminant extends the FD by mapping the data into a higher dimensional feature space $\mathcal{H}$.

In practice the KFD can be calculated implicitly by substituting all dot products with a kernel function $\kappa(\cdot, \cdot)$. It can be shown, that all positive definite kernel functions correspond to a dot product after transforming the data to a feature space $\mathcal{H}$ with mapping $\boldsymbol{\Phi}(\cdot)$ (Schölkopf & Smola, 2001). The feature space $\mathcal{H}$ can be very high dimensional, and the use of the projection vector $\mathbf{w}$ directly may be impractical or impossible. Instead, the famous *Representer theorem* guarantees that the solution can be represented as a linear combination of the mapped samples (Schölkopf & Smola, 2001). Thus, the Fisher ratio in the feature space is based on the weights of the samples $\boldsymbol{\alpha}$ instead of the weights of the dimensions:

$$J(\boldsymbol{\alpha}) = \frac{\boldsymbol{\alpha}^T \mathbf{Q}^T \mathbf{S}_B^{\boldsymbol{\Phi}} \mathbf{Q} \boldsymbol{\alpha}}{\boldsymbol{\alpha}^T \mathbf{Q}^T \mathbf{S}_W^{\boldsymbol{\Phi}} \mathbf{Q} \boldsymbol{\alpha}}, \tag{5}$$

where $\boldsymbol{\alpha} \in \mathbf{R}^N = (\alpha_1, \alpha_2, \ldots, \alpha_N)^T$ is the weight vector for the mapped training samples in the matrix $\mathbf{Q} = [\boldsymbol{\Phi}(\mathbf{x}_1), \ldots, \boldsymbol{\Phi}(\mathbf{x}_N)]$, and $\mathbf{S}_B^{\boldsymbol{\Phi}}$ and $\mathbf{S}_W^{\boldsymbol{\Phi}}$ are the between-class and within-class scatter matrices in the feature space $\mathcal{H}$, respectively.

Similar solution as the one for the Fisher discriminant in Eq. (4) can be found also for this case (Schölkopf & Smola, 2001). However the inversion becomes more difficult, since the dimension of the weight vector $\boldsymbol{\alpha}$ is now the number of the collected training samples. Therefore, we need a regularization term $\lambda\mathbf{I}$, where λ is a small positive scalar and $\mathbf{I}$ is the $N \times N$ identity matrix. Using regularization also improves the robustness of the projection and makes it less likely to overfit, as the solution becomes less sensitive to within-class scatter. In our notation this yields the solution

$$\boldsymbol{\alpha} = (\mathbf{Q}^T \mathbf{S}_W^{\boldsymbol{\Phi}} \mathbf{Q} + \lambda\mathbf{I})^{-1} \mathbf{Q}^T (\boldsymbol{\mu}_1^{\boldsymbol{\Phi}} - \boldsymbol{\mu}_0^{\boldsymbol{\Phi}}), \tag{6}$$

where $\boldsymbol{\mu}_1^{\boldsymbol{\Phi}} \in \mathcal{H}$ and $\boldsymbol{\mu}_0^{\boldsymbol{\Phi}} \in \mathcal{H}$ are the sample means of the mapped wheal and skin samples, respectively.

It is straightforward to show, that Eq. (6) can be expressed in terms of dot products and thus the kernel trick (Mika et al., 1999; Schölkopf & Smola, 2001). Also the actual projection of a test sample $\mathbf{x} \in \mathbf{R}^3$ can be expressed through the kernel as

$$y = \boldsymbol{\alpha}^T \mathbf{Q}^T \boldsymbol{\Phi}(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i \kappa(\mathbf{x}_i, \mathbf{x}). \tag{7}$$

To reveal the kernel nature of (6), a part of it is further investigated, giving

$$\mathbf{Q}^T\mathbf{S}_W^{\boldsymbol{\Phi}}\mathbf{Q} = \sum_{i=0,1} \sum_{\mathbf{x}\in\mathcal{S}_i} \mathbf{Q}^T(\boldsymbol{\Phi}(\mathbf{x}) - \boldsymbol{\mu}_i^{\boldsymbol{\Phi}})(\mathbf{Q}^T(\boldsymbol{\Phi}(\mathbf{x}) - \boldsymbol{\mu}_i^{\boldsymbol{\Phi}}))^T. \tag{8}$$

Now it is rather clear that (6) can be represented using only dot-products of the mapped samples. Since also the projection of a test sample $\mathbf{x}$ to grayscale is expressible as $\boldsymbol{\alpha}^T\mathbf{Q}^T\boldsymbol{\Phi}(\mathbf{x})$, the explicit use of mapping function $\boldsymbol{\Phi}(\mathbf{x})$ can be substituted with its dot-products, which in hand are replaceable with Mercer kernels (Mika et al., 1999). This *kernel-trick* removes the curse of dimensionality, ultimately allowing implicit mapping to an infinite-dimensional space. More precise introduction, although with different notation, to KFD by Mika *et al.* is in Mika et al. (1999) and in Schölkopf & Smola (2001).

There are various alternatives for the kernel function $\kappa(\cdot,\cdot)$, among which the most widely used are the polynomial kernels and the Radial Basis Function (RBF) kernel. We experimented with various kernels, and found out that the polynomial kernels do not increase the separation significantly when compared with the linear kernel, which is equivalent to the traditional FD. In other words, all low-order polynomial kernels produce a projection very similar to the first order kernel, shown in Figure 6 (e). However, the separation seems to improve with the RBF kernel

$$\kappa(\mathbf{u},\mathbf{v}) = \exp\left(-\frac{||\mathbf{u}-\mathbf{v}||^2}{2\sigma^2}\right). \tag{9}$$

There are two parameters in the KFD projection with RBF kernel: The regularization parameter λ and the kernel width σ^2. Since there exists a lot of training data in our case, it seems to be less sensitive to the regularization parameter λ than the width σ^2. In our experiments we set the value of $\lambda = 10^{-5}$, and if the condition number of the matrix in Eq. (6) indicates that the matrix is close to singular, the value of λ is increased ten-fold until the inversion succeeds.

Figures 6 (f-g) illustrate the effect of the bandwidth parameter σ^2. Figure 6 (f) uses the bandwidth selected using the so called Silverman's rule of thumb (Silverman, 1986), widely used in kernel density estimation and defined by $\hat{\sigma}_{\mathrm{rot}} = 1.06\hat{\sigma}_x N^{-\frac{1}{5}}$, where $\hat{\sigma_x}$ is the sample standard deviation of the data and N is the data length. In the example in Figure 6 (f) the rule of thumb gives $\sigma_{\mathrm{rot}} = 1.37$. Figure 6 (g) on the other hand illustrates the result with fixed $\sigma = 0.5$.

It seems that the rule of thumb tends to give too large values for our problem. This can be seen from the visually improved separation which is obtained using a smaller bandwidth, $\sigma = 0.5$, whose result is shown in Figure 6 (g).

3.5 Segmentation of the grayscale image

As the final step of the wheal area segmentation, one has to threshold the grayscale projected image in order to obtain a binary segmentation. The simple approach is to use thresholding. The problem of selecting a proper threshold has been extensively studied, and, e.g., Sezgin *et al.* (Sezgin & Sankur, 2004) compares 40 selected methods for a set of test images. One of their conclusions is that the best performing method is different depending on the nature of the input grayscale image (e.g., text, natural scenes, etc.). Since one of our goals is to compare different grayscale projections for the wheal detection purpose, we restrict ourselves
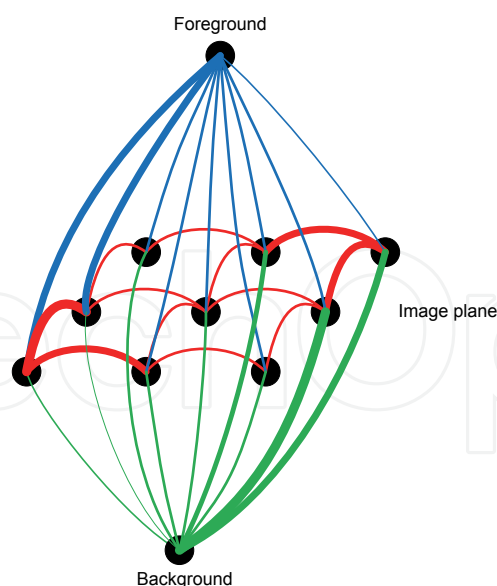
Fig. 7. The graph representation of a 3×3 image plane.

to using only one automatic threshold selection method; the most widely used Otsu method (Otsu, 1979). Without this restriction, we would be comparing all combinations of grayscale projection and thresholding techniques summing up to hundreds of combinations.

As an alternative to grayscale thresholding, we consider a newer approach, the powerful graph cut based segmentation. The idea of using graphs for solving hard optimization problems was originally discovered by Greig *et al.* (Greig et al., 1989), and extended to multilabel problems by Boykov *et al.* (Boykov et al., 2001). A fast implementation was proposed by Boykov *et al.* (Boykov & Kolmogorov, 2004), and their free implementation is widely used, and is the basis for our method, as well.

Graph cuts are a method for efficient minimization of certain energy functionals of the type

$$E[f] = \sum_{p \in \mathcal{P}} D_p(I_p) + \sum_{p,q \in \mathcal{N}} V_{p,q}(I_p, I_q), \tag{10}$$

where $D_p(f_p)$ reflects the distance from the image to be modeled (segmented in our case), and $V_{p,q}(I_p, I_q)$ is the cost assigned to labeling neighboring points differently. In a sense, using graph cuts is *regularized thresholding*, where neighboring points are encouraged to have the same label. In practice, the difference to thresholding is that the resulting labeling has fewer holes. In our case, the graph cut approach for segmentation treats the KFDA-projected grayscale image as a weighted graph, whose nodes represents the pixels and edges the connections between the neighboring pixels. An illustration of this is shown in Figure 7, which represents a 3×3 image, whose nine pixels are considered as nodes of the graph. Each pixel is connected to its closest neighbors, and additionally to two special nodes; *foreground* node and *background* node.

The graph cut method attempts to split the graph into two disconnected parts, such that the foreground and background nodes are in different partitions. The splitting is done with minimal cost, i.e., such that the sum of weights of cut nodes is as small as possible. As a result, the foreground area consists of pixels connected to the foreground node after splitting

the graph. Note, that in our case an alternative graph formulation could take advantage of the known foreground and background locations: the background node could be connected only to the borders of the image (which is known to be background) and the foreground node only to the center of the image.

In order to obtain a reasonable segmentation result, the weights are determined according to the following rules. Edges connecting neighboring pixels have a weight inversely proportional to the difference of the grayscale values. This way homogeneous areas (with grayscales close to each other) have large weights and heterogeneous areas (large pixel difference) have small weights. In our case the weight of the edge between pixels p and q with grayscale intensities I_p and I_q is determined by the rule

$$V_{p,q}(I_p, I_q) = 10 \cdot \exp\left(-\frac{|I_p - I_q|}{100}\right). \tag{11}$$

This function satisfies the requirement that far away grayscale values have a small weight while close grayscales obtain a large weight. The exact exponential form and the normalization coefficients of the function were obtained through experimentation, although various other choices were almost equally effective.

The edges connecting the pixels to *foreground* and *background* nodes are determined by the grayscale value. The idea is that the brighter the pixel, the stronger the connection to the foreground node and vice versa. In our case the foreground edge weight for pixel p with intensity $I_p \in \{0, 1, \ldots, 255\}$ was determined by the rule

$$D_f(I_p) = \frac{\gamma}{255 - I_p + 1}, \tag{12}$$

and the background edge weight by the rule

$$D_b(I_p) = \frac{1 - \gamma}{I_p + 1}. \tag{13}$$

The idea is that the foreground connection would be strong, when the pixel value I_p is large (close to 255) and the background connection would be strong, when the pixel value I_p is small (close to 0).

The parameter $\gamma \in [0, 1]$ balances the edge weights and can be used for adjusting the foreground area size. There are closed form solutions for the parameter to obtain a desired the probability of a foreground pixel. However, an easier method is to use binary search over $\gamma \in [0, 1]$ to get a desired ratio of background and foreground sizes. In our case we selected the desired ratio to be equal to that of the Otsu thresholding result.

As the final step, we remove all but the largest object from the segmentation result. This is because sometimes there remain small foreground areas that are due to noise.

3.6 Using a shape model for wheal area detection

The transition from the background (the healthy skin) to the foreground (the wheal) can be quite smooth, and the KFD-projected image may contain several individual foreground
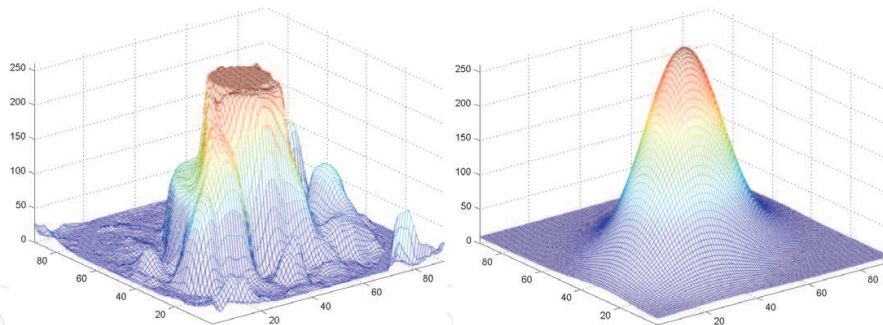
Fig. 8. Left: The grayscale projection data of the wheal in Figure 6 using KFD projection with RBF kernel and $\sigma = 1$. Right: The final result after nonlinear LS fit.

regions although the image has only one wheal. This is mostly due to the noise in the data, whose effect is greatly emphasized by the grayscale projection.

In order to increase the robustness of the segmentation, we fit a shape model for the appearance of the wheal. Since the manual measurement assumes that the wheals are ellipses, an elliptic shape model seems reasonable. Thus, the problem is to find an ellipse that divides the image into two maximally inhomogeneous areas.

Since there are an infinite amount of ellipses, we have to limit the search space somehow. This can be done by fitting a model to the grayscale projection and considering only the isosurfaces of the model. Based on Figure 6, the Gaussian surface seems an appropriate model for the spatial grayscale distribution in this case. Moreover, it suits our assumption of elliptic wheals, because the isosurfaces of the two-dimensional Gaussian are ellipses.

More specifically, the Gaussian model is defined by

$$f(\mathbf{x}; c, \mathbf{x}_0, \mathbf{\Sigma}) = c \cdot \exp\left(-(\mathbf{x} - \mathbf{x}_0)^T \mathbf{\Sigma} (\mathbf{x} - \mathbf{x}_0)\right), \tag{14}$$

where $c \in \mathbf{R}_+$ defines the scale of the Gaussian, $\mathbf{x} = (x, y)^T$ denotes the image coordinates where the the model is fitted, $\mathbf{x}_0 = (x_0, y_0)^T$ denotes the location of the peak of the Gaussian and $\mathbf{\Sigma} \in \mathbf{R}^{2 \times 2}$ is a symmetric coefficient matrix.

The least squares (LS) fit to the grayscale image data is defined by

$$\min_{c, \mathbf{\Sigma}, \mathbf{x}_0} \sum_{k=0}^{N} (z_k - f(\mathbf{x}_k; c, \mathbf{x}_0, \mathbf{\Sigma}))^2, \tag{15}$$

where z_k denotes the grayscale value at image position $\mathbf{x}_k$. Note that the data has to be preprocessed by subtracting the minimum of z_k, $k = 0, \ldots, N$, in order to avoid a constant offset term in the model.

Fitting the Gaussian is a nontrivial problem, although lot of literature on the topic exists (e.g., Brändle et al. (2000)). However, the easiest approach is to use software packages such as Matlab Optimization toolbox to find the optimal parameters.
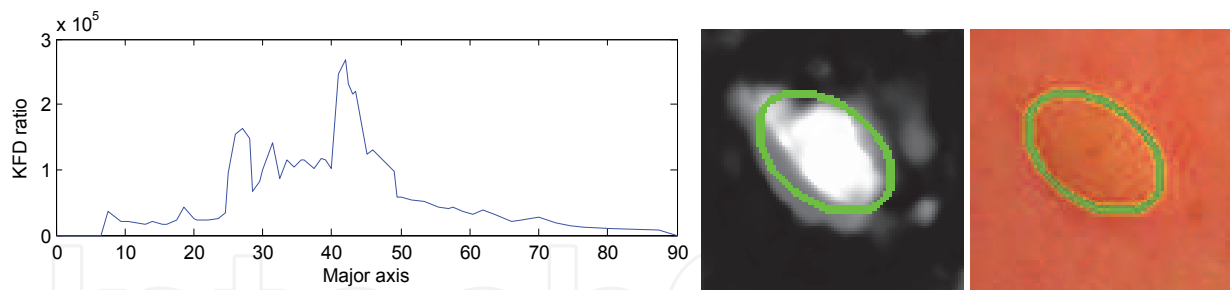
Fig. 9. Left: The KFDR as a function of the ellipse size. Center: The maximally separating ellipse overlaid on top of the corresponding KFD projection. Right: The maximally separating ellipse overlaid on top of original RGB data.

However, in order to ease the task, we first seek for an initial guess for the coefficients by taking the logarithm of the data and the model of Eq. (15):

$$\min_{c,\mathbf{\Sigma},\mathbf{x}_0} \sum_{k=0}^{N} (\ln(z_k) - \ln(f(\mathbf{x}_k; c, \mathbf{x}_0, \mathbf{\Sigma})))^2 = \min_{c,\mathbf{\Sigma},\mathbf{x}_0} \sum_{k=0}^{N} (\ln(z_k) + \ln c + (\mathbf{x} - \mathbf{x}_0)^T \mathbf{\Sigma}(\mathbf{x} - \mathbf{x}_0))^2. \quad (16)$$

This makes the problem linear, and the result can be found in a closed form. However, taking the logarithm increases the importance of the smaller values, and the method essentially fits the model to the noise surrounding the wheal, not the wheal itself. Thus, the resulting solution is used only as the initial guess for the nonlinear LS problem without the logarithm. The nonlinear LS problem is then solved using Matlab Optimization toolbox.

Figure 8 shows the original grayscale data on the left, the result of logarithmic fitting in the center and the result of nonlinear iterative fitting on the right.

The isosurfaces of the Gaussian fit can be used as candidates for elliptic segmentation. As noted earlier, all the isosurfaces are cross-sections of a paraboloid and thus ellipses. Moreover, due to fitting, they most likely have the correct orientation and correct ratio of major and minor axis lengths. Thus, our next goal is to seek for the best elliptic isosurface among them all.

The definition of a good ellipse among the candidates needs some measure of separation between the segmented areas. Recent work by Harchaoui *et al.* (Harchaoui et al., 2008) considers using the Kernel Fisher Discriminant Ratio (KFDR) for testing the homogeneity between two sets, which coincides well with our use of KFD for grayscale projection in Section 3.2.

In other words, we test all ellipses that are cross sections of the fitted Gaussian and attempt to maximize the KFDR of Eq. (5) with respect to training sets defined by the ellipse. The situation is similar to the grayscale projection, but now we are not looking for a good classifier for the RGB data, but only assessing how well the data *could be classified*. Unlike Section 3.2, the choice of the training samples is now based on the boundaries of the ellipse to be tested.

In a sense, the KFDR homogeneity test attempts to design a classifier to separate the "inside" class and the "outside" class, and the KFDR is a natural measure how well this can be done. Note that this is not equivalent to calculating the variances directly from the projections of
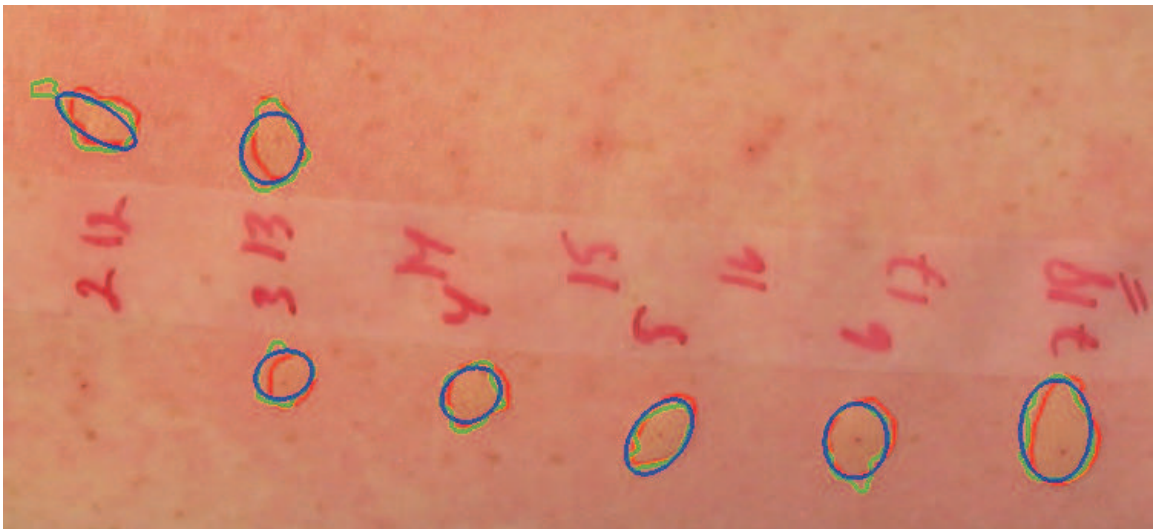
Fig. 10. An example of segmentation result. The red boundary is the result of manual segmentation, while blue and green boundaries represent the result of our method with and without the elliptic shape model, respectively. (Wheals 2, 14, 15, etc. have not been segmented, as there is no detectable allergic reaction at these sites.)

Figure 6, because the projection is calculated separately for the training sets determined by each ellipse candidate.

Sometimes the KFDR separability criterion results in very small ellipses, because a small foreground training set tends to be well separable. As an extreme example, an ellipse containing only a single pixel has extremely good separability assuming no other pixel has exactly the same RGB value. Thus, we decided to modify the criterion by multiplying it with the cardinality of the smaller training set. Alternatively, we could set a minimum size restriction for the ellipse.

An example of the separability test is shown in Figure 9. The figure shows the KFDR between the "inside" class and the "outside" class for ellipses with different radius. It can be seen that the maximal separation is obtained at radius 42, and the corresponding ellipse is illustrated in Figure 9, as well.

3.7 Experiments

The results from the described method are compared to manual wheal segmentations (made by a non-medical expert). The similarity measure used by Celebi et al. (Celebi et al., 2009) compares the *areas* of the segmentations. For our purposes, this is not an appropriate criterion, since ultimately we are interested in the major and minor axes of the wheal. The error in areas increases quadratically with respect to the axes, which is not desirable. Instead, we used the following error criterion between the computer segmentation A and the manual ground truth B:

$$E(A,B) = \frac{\sqrt{\text{Area}(\text{OR}(A,B))} - \sqrt{\text{Area}(\text{AND}(A,B))}}{\sqrt{\text{Area}(B)}}, \tag{17}$$

where $\text{OR}(A,B)$ consists of pixels segmented as foreground in A or B, and $\text{AND}(A,B)$ of foreground pixels in both A and B. Moreover, $\text{Area}(A)$ is the number of foreground pixels

in A. The favourable property of Eq. (17) is that it increases linearly with respect to the error in major and minor axes. For example, it can be shown that the error measure for concentric circles with radii $r + a$ and $r - a$ are equal if the true radius is r. This is not the case with the error of (Celebi et al., 2009).

Examples of segmentation results are illustrated in Figure 10. The figure shows the result of manual segmentation (red) compared with the result of the proposed method with (blue) and without (green) the shape model. Table 2 represents the average errors with different grayscale transformations. The test data consists of seven wheals including those shown in Figure 10. The five first columns represent different KFD projections designed using training data, while in the last two columns the projection is designed in an unsupervised or *ad hoc* manner.

From the results, one can see that the $L^*a^*b^*$ color space produces the smallest errors. This is in coherence with the results of earlier studies (Nischik & Forster, 1997; Roullot et al., 2005), who have also discovered the importance of the a^* component in skin color analysis. However, the difference between color spaces is not that significant, especially when using a Gaussian kernel for the KFD. For example, in the case of Gaussian kernel with bandwidth $\sigma = 1$, all errors are close to each other, and visual inspection of the results reveals that there are no gross errors. The lesser importance of the initial color space when using the Gaussian KFD kernel is natural, because the three-dimensional data is mapped to an infinite dimensional space, where the colors are most likely separable regardless of the initial color space.

Another interesting observation is the worse than expected performance of the elliptic shape model. In many cases its use increases the error. However, this is partly due to the fact that the manually segmented wheals are not ellipses, so the elliptic model can not reach zero error even in theory. The best cases are the ones where the true wheal is ellipse-shaped with no elongated pseudopodia, e.g., the 2$^{\text{nd}}$ and 3$^{\text{rd}}$ wheals from the right in Figure 10. In all other cases the wheal shape is more irregular, and the shape model results in the largest inscribed ellipse. However, there is some randomness in the results due to the small N. We plan to

			Gaussian ($\sigma = 0.5$)	Gaussian ($\sigma = 1$)	Gaussian ($\sigma = \sigma_{\text{rot}}$)	Linear kernel	2. order kernel	Celebi method	Ad hoc
With Otsu Thresh		RGB	0.3668	0.2393	0.1912	0.2567	0.2723	1.3710	1.0297
		$L^*a^*b^*$	0.2505	0.1861	0.2811	0.2382	0.2925	1.0788	0.9835
		HSV	0.2177	0.2176	0.5129	0.2733	0.2563	1.5259	0.7424
With Graph Cuts		RGB	0.3660	0.2406	0.1915	0.2545	0.2714	1.3759	1.0307
		$L^*a^*b^*$	0.2508	0.1855	0.2813	0.2364	0.2926	1.0810	0.9844
		HSV	0.2176	0.2169	0.5129	0.2714	0.2570	1.5260	0.7422
With Shape Model		RGB	0.4769	0.2479	0.2289	0.4821	0.2143	1.3588	0.4291
		$L^*a^*b^*$	0.2325	0.2388	0.2335	0.1894	0.2240	1.1731	1.4467
		HSV	0.2472	0.2300	0.7174	0.4852	0.3815	1.3465	0.8816

Table 2. The comparison of automated wheal measurement methods in terms of the error of Eq. (17). Each row defines an initial color space, each column corresponds to a grayscale transformation. The last column corresponds to a manually designed transformation based on what looks good. In the RGB case, the *Ad Hoc* transformation is the difference $G - B$, in the $L^*a^*b^*$ case it is the a^* component, and in the HSV case the H component.

study the performance with larger *N* and compare them with the manual results of a trained physician.

4. Conclusions

This research projects shows one possibility to automatize a simple dermatological examination. While the results obtained are not perfect, they are very promising. Automatization gives several benefits in the case of the prick test, as it eliminates the inter-observer variance and offers a way to study the immune reaction as a function of time.

The method shown above is semi-automatic in the sense that the user has to input a point close to the center of the wheal. This is not a major obstacle in practice, as the prick test can be performed so that the centers are known. Another possibility is to use some easily recognizable markings on the skin so that the centers can be found automatically.

Finding the size of a wheal is a surprisingly complicated task. The wheal edges are not sharp, and the contrast between the wheal and its surroundings is very low. The skin color is not constant, it varies significantly not only between individuals but also between different areas of the skin of an individual.

Our algorithm for determining the wheal size has several steps. First, the image is transformed into a grayscale image by using the KFD (Kernel Fisher Discriminant) projection. Although the linear kernel with the $L^*a^*b^*$ color space resulted in best results in this study, we anticipate that the flexibility of the Gaussian kernel would be useful with larger amount of patients (e.g., with different skin colors).

At this point there are two different methods for finding out the wheal size and shape. The first method segments the wheals using either straightforward thresholding or the Graph Cut method to determine which pixels in the image belong to the wheal. It seems that the performance of the segmentation approaches is equally good. The second method finds the ellipse which has the highest KFDR (Kernel Fisher Discriminant Ratio) between the pixels inside and outside of the ellipse.

It is not clear which of the methods is best in practice. The Graph Cut method gives irregular wheals, and the results are close to those obtained with manual segmentation. On the other hand, the medical practice uses one or two diameters of the wheal, and it in unclear whether or not any protruding features of the wheal should be included. Choosing between these two methods will require a large number of images, as the differences are not very large.

The images used in developing the algorithms were not of the highest quality. They were taken with a relatively old digital camera, and we earned during the process that there were significant compression artefacts in the images, which is clearly visible in the grayscale images. The use of a more modern camera without compression should reduce both compression artefacts and image noise significantly and thus improve segmentation results. This, also, is a topic for further research.

One should also note that the above methods generalize to other multichannel segmentation methods than RGB segmentation. Using alternative wavelengths might help the segmentation, and any combination of input images is possible. Another alternative is to generate artificial input channels, e.g., by filtering the RGB channels using different

preprocessing methods. As the results show, there is a difference in results when using different color spaces as the input. Thus, one might anticipate that adding artificial channels produced by different filters may also be of help in the segmentation.

5. References

Adobe Systems Incorporated (2009). *Digital Negative Specification version 1.3.0.0.*

Boykov, Y. & Kolmogorov, V. (2004). An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26(9): 1124 –1137.

Boykov, Y., Veksler, O. & Zabih, R. (2001). Fast approximate energy minimization via graph cuts, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 23(11): 1222 –1239.

Brändle, N., Chen, H., Bischof, H. & Lapp, H. (2000). Robust parametric and semi-parametric spot fitting for spot array images, *ISMB-2000 8th Intl. Conf. on Intell. Syst. for Mol. Biol.*, pp. 1–12.

Celebi, M., Iyatomi, H. & Schaefer, G. (2009). Contrast enhancement in dermoscopy images by maximizing a histogram bimodality measure, *16th IEEE Int. Conf. on Image Proc. (ICIP)*, pp. 2601–2604.

Coffin, D. (2011). dcraw source code.
    URL: *http://www.cybercom.net/ dcoffin/dcraw/dcraw.c*

ECMA (2009). *Technical Report TR/98 JPEG File Interchange Format (JFIF)*.

Farrell, J., Xiao, F. & Kavusi, S. (2006). Resolution and light sensitivity tradeoff with pixel size, *SPIE Electronic Imaging '2006 Conference*.

Fisher, R. (1936). The use of multiple measurements in taxonomic problems, *Annals of Eugenics* 7: 179–188.

Gomez, D., Butakoff, C., Ersboll, B. & Stoecker, W. (2008). Independent histogram pursuit for segmentation of skin lesions, *IEEE Trans. Biomed. Eng.* 55(1): 157–161.

Gomez, D. D., Clemmensen, L. H., ll, B. K. E. & Carstensen, J. M. (2007). Precise acquisition and unsupervised segmentation of multi-spectral images, *Comp. Vis. and Image Understanding* 106(2-3): 183–193.

Greig, D. M., Porteous, B. T. & Seheult, A. H. (1989). Exact maximum a posteriori estimation for binary images, *Journal of the Royal Statistical Society.* 51(2): pp. 271–279.

Haahtela, T., Petman, L., Järvenpaa, S. & Kautiainen, H. (2010). [Quality to allegy testing and interpretation of its results], *Duodecim* 126: 529–535.

Harchaoui, Z., Bach, F. & Eric, M. (2008). Testing for homogeneity with kernel fisher discriminant analysis, *Adv. in Neural Inf. Proc. Syst. 20*, MIT Press, Cambridge, MA, pp. 609–616.

Hecht, E. (1987). *Optics*, 2nd edn, Addison-Wesley, chapter 10.2.6.

Huttunen, H., Ryynänen, J.-P., Forsvik, H., Voipio, V. & Kikuchi, H. (2011). Kernel fisher discriminant and elliptic shape model for automatic measurement of allergic reactions, *in* A. Heyden & F. Kahl (eds), *Image Analysis*, Vol. 6688 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, pp. 764–773.

Meinert, R., Frischer, T., Karmaus, W. & Kuehr, J. (1994). Influence of skin prick test criteria on estimation of prevalence and incidence of allergic sensitization in children, *Allergy* 49: 526–532.

Mika, S., Ratsch, G., Weston, J., Scholkopf, B. & Mullers, K. (1999). Fisher discriminant analysis with kernels, *Proc. IEEE Neural Netw. for Signal Process. IX*, pp. 41–48.

NEMA (2011). *Digital Imaging and Communications in Medicine (DICOM)*.

Nischik, M. & Forster, C. (1997). Analysis of skin erythema using true-color images, *IEEE Trans. Med. Imag.* 16(6): 711–716.

Oppenheimer, J., Durham, S. & Nelson, H. (2011). Allergy diagnostic testing.
       URL: *http://www.worldallergy.org/professional/allergic_diseases_center/allergy_diagnostic/*

Otsu, N. (1979). A threshold selection method from gray-level histograms, *Systems, Man and Cybernetics, IEEE Transactions on* 9(1): 62–66.

Roullot, E., Autegarden, J.-E., Devriendt, P. & Leynadier, F. (2005). Segmentation of erythema from skin photographs for assisted diagnosis in allergology, *Pattern Rec. and Image Anal.*, Vol. 3687 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, pp. 754–763.

Santos, R., Mlynek, A., Lima, H., Martus, P. & M, M. (2008). Beyond flat wheals: validation of a three-dimensional imaging technology that will improve skin allergy research, *Clin. Exp. Dermatol.* 33(6): 772–775.

Schölkopf, B. & Smola, A. J. (2001). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, 1st edn, The MIT Press.

Seibert, S. M., King, T. S., Kline, D., Mende, C. & Craig, T. (2011). Reliability of skin test results when read at different time points, *Allergy Asthma Proc* 32: 203–205.

Sezgin, M. & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation, *Journal of Electronic Imaging* 13(1): 146–168.

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman-Hall.

Viggiano, J. A. S. (2004). Comparison of the accuracy of different white balancing algorithms as quantified by their color constancy, *Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications V: Proceedings of the SPIE, volume 5301.*

Wöhrl, S., Vigl, K., Binder, M., Stingl, G. & Prinz, M. (2006). Automated measurement of skin prick tests: an advance towards exact calculation of wheal size, *Experimental Dermatology* 15(2): 119–124.
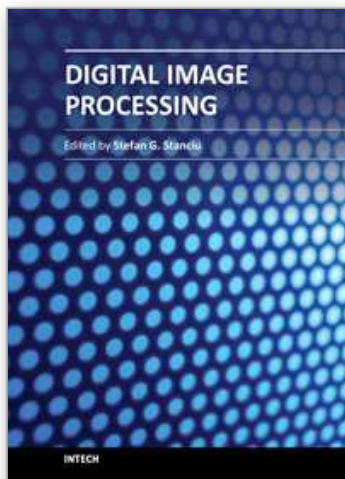
Digital Image Processing

Edited by Dr. Stefan G. Stanciu

This book presents several recent advances that are related or fall under the umbrella of 'digital image processing', with the purpose of providing an insight into the possibilities offered by digital image processing algorithms in various fields. The presented mathematical algorithms are accompanied by graphical representations and illustrative examples for an enhanced readability. The chapters are written in a manner that allows even a reader with basic experience and knowledge in the digital image processing field to properly understand the presented algorithms. Concurrently, the structure of the information in this book is such that fellow scientists will be able to use it to push the development of the presented subjects even further.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ville Voipio, Heikki Huttunen and Heikki Forsvik (2012). Practical Imaging in Dermatology, Digital Image Processing, Dr. Stefan G. Stanciu (Ed.), ISBN: 978-953-307-801-4, InTech, Available from: http://www.intechopen.com/books/digital-image-processing/practical-imaging-in-dermatology

INTECH
open science | open minds