

PRIP-TR-115  
Automatic Assessment of Lesion  
Development in Hemangioma  
Follow-Up Images

*Sebastian Zambanini*

**P**attern  
**R**ecognition &  
**I**mage  
**P**rocessing  
**G**roup



Institute of  
Computer Aided Automation

PRIP-TR-115

September 20, 2007

## Automatic Assessment of Lesion Development in Hemangioma Follow-Up Images

*Sebastian Zambanini*

### **Abstract**

This thesis presents an automatic method for the assessment of the development of cutaneous hemangiomas in digital images. The overall method provides two measurements on photographs taken during follow-up examinations: (1) the current skin area affected by the lesion and (2) the percentage/area of the hemangioma showing a regression, a so-called *graying*. For both analyses a pixel-wise classification scheme is applied to the images. The actual area measurement is accomplished through an image scale computation by means of a ruler attached to skin and visible in the images. Image registration is included in the assessment procedure to align follow-up images providing a direct comparison of color values necessary for a reliable detection of regressions. For image registration a robust feature-based method is presented that is able to deal with changing hemangioma appearances during follow-up. Experimental results are reported for the individual algorithms presented as well as for the whole procedure applied to follow-up images. In general, the proposed procedure working on follow-up images is expected to provide a more accurate and objective evaluation of the course of disease than the current clinical practice of manual measurement during an examination.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Cutaneous Hemangiomas . . . . .	1
1.2	Motivation . . . . .	2
1.3	Data Description . . . . .	3
1.4	Related Work . . . . .	3
1.5	Workflow . . . . .	5
1.6	Main Contribution . . . . .	6
1.7	Thesis Structure . . . . .	6
<b>2</b>	<b>Measurement of Hemangioma Size</b>	<b>7</b>
2.1	Image Scale Computation . . . . .	7
2.2	Segmentation of Hemangiomas . . . . .	8
2.2.1	Evaluation of Methods for Skin Lesion Segmentation . . . . .	9
2.2.2	Preprocessing . . . . .	12
2.2.3	Pixel-wise Classification . . . . .	15
<b>3</b>	<b>Registration of Follow-Up Hemangioma Images</b>	<b>22</b>
3.1	Introduction to Image Registration . . . . .	22
3.1.1	Feature-Based Methods . . . . .	23
3.1.2	Area-Based Methods . . . . .	24
3.1.3	Methods used for Skin Lesion Registration . . . . .	24
3.2	Robust Feature-Based Method for Registration of Follow-Up Hemangioma Images . . . . .	24
3.2.1	Detection of Interest Points in Hemangioma Images . . . . .	26
3.2.2	Matching of Interest Points . . . . .	30
3.2.3	Transformation Estimation Based on Interest Point Matches . . . . .	32
<b>4</b>	<b>Detection of Regressing Hemangioma Regions</b>	<b>37</b>
4.1	Pixel-wise Classification . . . . .	39
4.1.1	Neural Networks . . . . .	39
4.1.2	Feature Selection . . . . .	41
4.2	Regression Detection in Registered Hemangioma Follow-up Images . . . . .	42
<b>5</b>	<b>Experiments</b>	<b>44</b>
5.1	Measurement of Hemangioma Size . . . . .	44

5.1.1	Accuracy of the Scale Computing Algorithm . . . . .	44
5.1.2	Accuracy of Hemangioma Area Measurement on Synthetic Images .	46
5.1.3	Accuracy of the Hemangioma Segmentation Algorithm . . . . .	47
5.1.4	Precision of Automatic Measurement of Hemangioma Size . . . . .	48
5.2	Registration of Follow-Up Hemangioma Images . . . . .	49
5.2.1	Comparison of Difference-of-Gaussians and Canny Edge Interest Points . . . . .	50
5.2.2	Accuracy of Registration on Hemangioma Images taken at the same Time . . . . .	51
5.2.3	Accuracy of Registration on Follow-up Images . . . . .	52
5.3	Detection of Regressing Hemangioma Regions . . . . .	57
5.3.1	Accuracy of Regression Detection on Synthetic Follow-up Images .	57
5.3.2	Accuracy of Regression Detection on Real Follow-up Images . . . .	58
5.3.3	Accuracy of Regression Detection on Follow-up Images without Registration . . . . .	61
5.3.4	Precision of Regression Detection on Hemangioma Images taken at the same Time . . . . .	61
5.4	Discussion . . . . .	61
<b>6</b>	<b>Conclusion</b>	<b>65</b>
	<b>Bibliography</b>	<b>68</b>

# Chapter 1

## Introduction

The aim of the work presented in this master thesis is to provide dermatologists with a tool for assisting them in the assessment of skin lesion development, in particular the development of cutaneous hemangiomas. Since up to now the actual size and development was estimated through visual scoring by specialists by means of a small number of reference values [27] that are related to but do not measure the actual extent of lesions and affected areas, the assessment procedure can be made more accurate and objective by the use of computer vision methods on digital hemangioma images. Two analyses are made on the data:

1. The surveying of the area affected by the hemangioma by means of a ruler visible in the images.
2. The detection and surveying of regression (healing) regions.

The surveying of the area affected by the hemangioma is accomplished by image segmentation and determination of the image scale based on the ruler visible on the skin. In general, this procedure only surveys the clearly reddish, mostly unregressed regions of the hemangioma and can be applied to single images. In contrast, the detection and surveying of regressing regions describes a change over time, and therefore includes a comparison to the image data from the first examination, making an image registration of the individual images necessary to account for different view points.

In Section 1.1 a brief medical description of cutaneous hemangiomas is given, followed by a detailed motivation of the proposed work in Section 1.2. The provided clinical image data is described in Section 1.3. A brief overview of related work is given in Section 1.4. The overall workflow of the proposed procedure for assessment of hemangioma development is presented in Section 1.5. The main contribution of the thesis is stated in Section 1.6. Finally, Section 1.7 gives a structural overview of the thesis.

### 1.1 Cutaneous Hemangiomas

Cutaneous hemangiomas are benign tumors made up of newly-formed blood vessels in the skin and the most common benign vascular tumors in infancy with a frequency of

about 10 % [10]. Due to their potency for rapid proliferation they may threaten vital structures by tumor compression or tumor obstruction and/or may impair vital functions such as breathing, vision, hearing, ingestion or excretion [6]. Lesions occurring in the face or neck may cause psychological problems in the very young patients (typical age of patients is 6 - 24 months) and their parents. Safe and effective treatment modalities at the earliest time possible can stop further proliferation, induce regression and prevent complications [27]. In the past various different treatment methods have been described including steroids, cryotherapy, embolization, sclerotherapy, surgery and laser therapy [40]. During treatment, and during clinical trials comparing the performance of different treatment strategies, the accurate monitoring of the progress of hemangioma growth and regression is essential. More accurate measurements of the change provide better feedback during therapy, and allow for more significant results given a limited number of patients included in clinical trials.

## 1.2 Motivation

During the period of treatment of a hemangioma, besides the clinical examinations, photos of the hemangioma along with a ruler attached to the skin are taken to record its current appearance and area. An example for such a series of follow-up images is shown in Figure 1.1, presenting also the characteristic form of healing during the treatment. At the first examination (Figure 1.1(a)), the majority of the hemangioma shows a typical reddish saturated color caused by the excess of blood vessels and almost no regression has started yet. Only a small regression region is indicated by a so-called *graying* in the center of the hemangioma. At the second examination 8 weeks later (Figure 1.1(b)), the regression covers almost the whole hemangioma. Further 5 weeks later (Figure 1.1(c)), the whole hemangioma can be classified as regressing.

**State of the art:** In current clinical practice, the present hemangioma area is estimated by a simple manual measurement of the height and width of the lesion during a clinical examination. The degree of regression is reported by a visual estimation of the fraction of the hemangioma showing a graying of the typical reddish hemangioma color. Evidently, this way of assessment is (1) inaccurate since only the diameters and not the shape of hemangioma are considered in its surveying and (2) subjective since the estimation of the regressing hemangioma fraction has the tendency to vary strongly between dermatologists.

**Contribution:** Therefore, the motivation of the thesis was to make the assessment of the hemangioma's course of disease more accurate and objective by applying computer vision methods on digital images of the form shown in Figure 1.1. Generally, the benefit of such an assessment is twofold. Firstly, it can be used in clinical studies evaluating the efficacy of various treatment modalities for hemangiomas. In this context, the significance of the results is increased by a more accurate and objective evaluation. Secondly, the

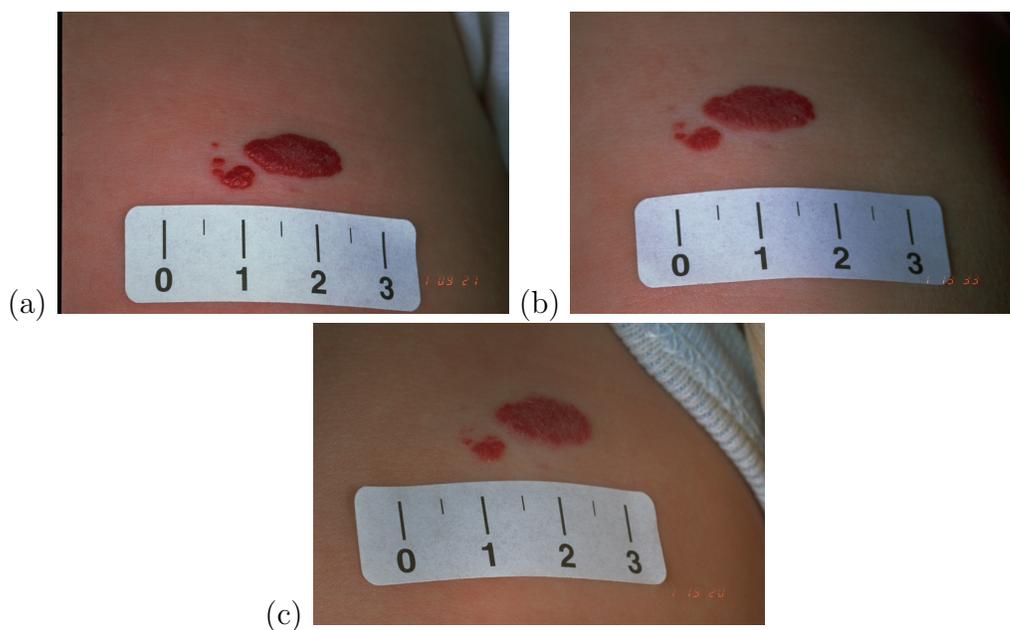


Figure 1.1: A follow-up series of hemangioma images.

assessment can be used during treatment to validate or invalidate the efficacy of the chosen treatment modality.

### 1.3 Data Description

The images used in this thesis have been kindly provided by Dr. Harald Maier from the *Division of Special and Environmental Dermatology* at the *Medical University of Vienna*. The photos were taken using an analog photo camera and digitized with a scanner. All images have a resolution of 512x768 pixels and a bit depth of 8 bits per color channel. In order to increase reliability there are always two or more images taken of the hemangioma, i.e. there are at least two images of a specific hemangioma at a specific examination date. All images are labeled with a five-digit code in the lower right corner to uniquely identify them.

Because of the circumstance that nearly all patients in the study are infants who normally do not keep still when being photographed and the bad illumination in the examination room, the quality of the images is partially poor. Two examples are shown in Figure 1.2. In Figure 1.2(a) parts of the hemangioma are situated in an inadequately illuminated region and Figure 1.2(b) is generally of low sharpness.

### 1.4 Related Work

This section gives a short survey about related work on the use of computer vision methods on digital images of skin lesions. An overview of the specific approaches for skin

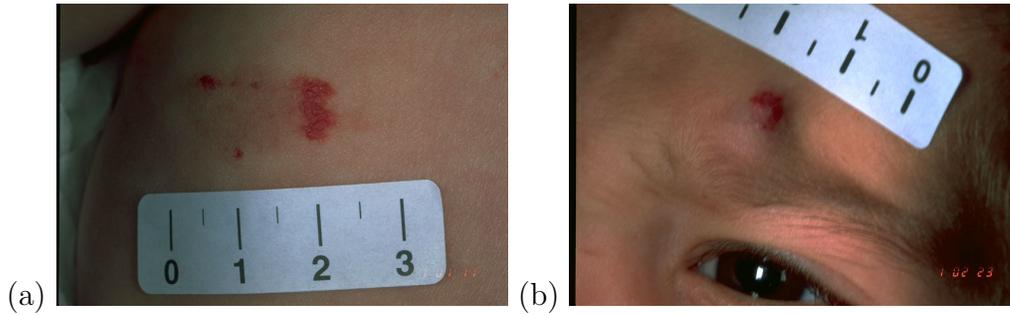


Figure 1.2: Poor quality hemangioma images.

lesion segmentation and registration are given in the respective chapters of the thesis, thus segmentation methods are discussed in Section 2.2.1 and registration methods for follow-up studies in Section 3.1.3.

To our knowledge, no research has been done before on the image-based assessment of cutaneous hemangiomas. Nonetheless, various other types of skin lesions were covered. Especially, the automatic image-based classification of melanoma in benign and malignant lesions has been studied extensively in literature starting as early as 1987 [4]. An overview of work published until the year 2000 is given in [8]. Recent approaches are summarized in [5]. Classification of melanoma is thereby typically based on the *ABCD* rule (**A**symmetry, **B**order irregularity, **C**olor variegation and **D**iameter) used by dermatologists [32].

Another skin lesion type covered is psoriasis [18]. In psoriasis assessment, a *psoriasis area and severity index* (PASI) [39] is used to evaluate the degree of disease. A variable thresholding technique for area computation as part of this assessment method is proposed by Roening et al. [34]. The automatic change detection in follow-up psoriasis images is addressed as well: Maletti and Ersbøll propose a method for change detection in both registered [28] and unregistered images [29], each one by applying the Multivariate Alteration Detection Transform (M.A.D.) on the data. In [14] two other methods for change detection are explored: simple image subtraction and Principle Component Analysis (PCA). A method for area computation using Active Contour Models (ACMs) is presented in [22] and applied on images of leg ulcers.

However, the described methods cannot be applied directly to cutaneous hemangiomas, because the sole measurement of affected skin area cannot be used as a complete assessment method for this lesion type, since the regression process starts with color changes (graying) in certain hemangioma regions and not with their complete disappearance. The change detection scheme proposed by Maletti and Ersbøll [28][29] shows suitable results on psoriasis images but is designed for the detection of texture changes, whereas changes in hemangioma images are indicated by differing color values during follow-up. Nonetheless, the method proposed in this thesis is considered to be applicable on other lesion types as well, although experiments were only conducted on hemangioma images. Both the segmentation of the affected skin area and the detection of regression regions are based on a pixel-wise classification by a previously trained classifier, thus the method could be adapted to other lesion types if an adequate number of annotated training data is available.

# 1.5 Workflow

This section describes the whole workflow for the assessment of hemangioma development. Additionally, some terms are defined here that will consistently be used in the remainder of the thesis.

The workflow is illustrated in Figure 1.3. Here four follow-up images of a specific hemangioma are given and the determination of the hemangioma area and the measurement of regions showing a regression in Image 4 is exemplarily shown. In Figure 1.3 the hemangioma region is indicated by a white area whereas the expanding regression during follow-up is indicated by a gray area. The following steps are conducted:

1. **Hemangioma Segmentation:** The image regions of Image 1 belonging to the hemangioma, called *hemangioma regions* are determined. In Figure 1.3, this segmentation is indicated by the black border around the white area.
2. **Registration and Transformation:** The transformation aligning Image 1 with Image 4 is determined to define a region of interest (ROI) corresponding to the expected location of the hemangioma in Image 4. This new region is again indicated by a black border. It is calculated based on the hemangioma segmentation in Image 1 and the transformation of Image 1 to Image 4.
3. **Regression Detection:** The regions showing a regression (the *regression regions*) are determined in the ROI, including a comparison of color values between Image 4 and the transformed Image 1.

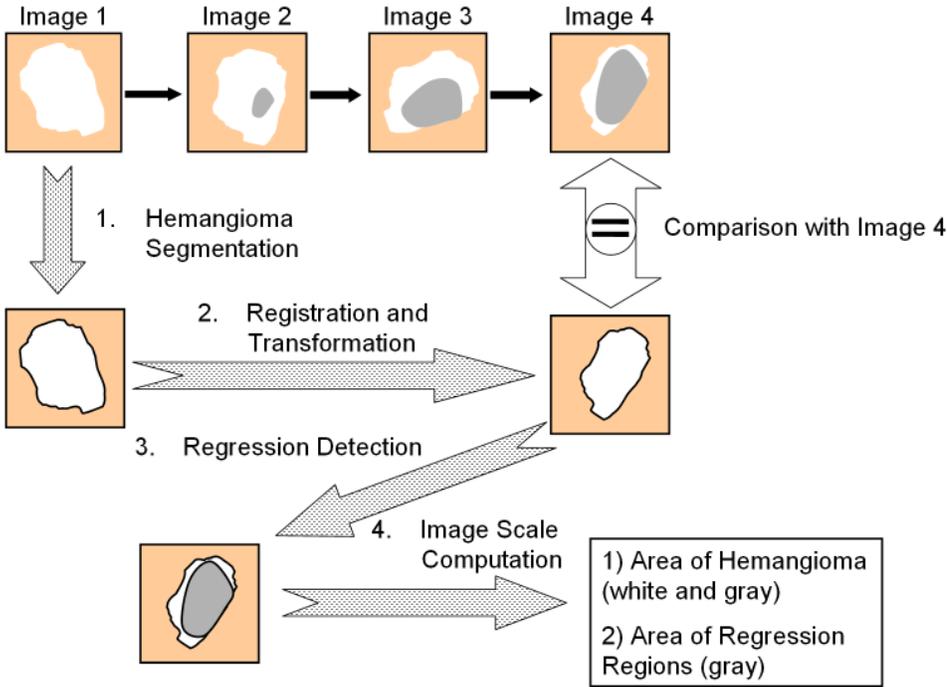


Figure 1.3: Illustration of the workflow for the assessment of hemangioma development.

4. **Image Scale Computation:** The ruler visible in Image 4 is used to determine the image scale. By combining the information of image scale and regression/hemangioma regions, the current area of both the hemangioma and the regression regions is obtained.

## 1.6 Main Contribution

The thesis proposes a method to automatically assess the development of cutaneous hemangiomas in follow-up images. Thereby, on the clinical side a more accurate assessment of medication effects can be achieved. This allows for a reliable evaluation of treatment methods, providing higher accuracy and consistency than the standard manual assessment. Furthermore, the method is of general interest in the field of automatic and image-based assessment of skin lesions: the use of a pixel-wise classification scheme on follow-up images of skin lesions is new and this concept makes the method highly adaptable to other kinds of lesions, e.g. melanoma, psoriasis or wounds.

## 1.7 Thesis Structure

The remainder of the thesis is organized as follows: The measurement of hemangioma size (i.e. step 1 and 4) is covered in Chapter 2. The method for registration of follow-up images (step 2) is treated in Chapter 3. Chapter 4 describes the method for regression detection (step 3). Experimental results of both the individual methods and the whole assessment procedure are reported and discussed in Chapter 5. A conclusion is given in Chapter 6.

Parts of this thesis were already published elsewhere: the segmentation and area measurement of the hemangioma region was presented in [42]. The registration method used for the alignment of follow-up images was presented in [43].

# Chapter 2

## Measurement of Hemangioma Size

This chapter deals with the area measurement of the region affected by a hemangioma on a single image. It achieves an independent measurement for a single acquisition time point and is also a necessary prerequisite for the detection of regressions occurring during follow-up examinations (Chapter 4). During image acquisition (e.g. during a clinical trial) all hemangiomas are photographed along with a ruler to determine the scale of the image. The problem can be divided into two main tasks:

1. Determining the scale of the images by means of the ruler visible in the image.
2. Segmentation of the skin area belonging to the hemangioma.

This chapter is organized as follows: Section 2.1 explains in detail the algorithm used for determining the scale of the images. In Section 2.2.1 first various segmentation algorithms are evaluated for our purpose, followed by a presentation of the complete segmentation process, divided into *Preprocessing* (Section 2.2.2) and *Pixel-wise Classification* (Section 2.2.3).

### 2.1 Image Scale Computation

All images show a ruler with 4 bold lines at 1 cm distance steps beside the hemangioma. Therefore, the task of the algorithm is to compute the Euclidean distance between two lines to obtain the spatial resolution of the images. The area of the hemangioma is then simply calculated by multiplying the number of hemangioma pixels by the area of one pixel. We neglect the error due to the fact that normally both hemangioma and ruler are not situated on a planar surface parallel to the image plane. Although this assumption of planarity is naturally violated by the shape of the hemangioma and skin, during image acquisition only small patches of skin (average hemangioma area is  $\sim 0.6 \text{ cm}^2$ ) are photographed and a maximum closeness between hemangioma and ruler as well as an optimal camera viewpoint is taken into consideration by the photographer. Thus, this kind of simplification is assumed to be a reasonable approximation. The main steps of the algorithm are:

1. Segment the ruler.

2. Calculate the ruler orientation and rotate it in a horizontal position.
3. Get the number of pixels between two lines by scanlining through the ruler.

**1. Segmenting the ruler:** First we have to segment the ruler. Since all rulers are white and differ extremely from the rest of the image, this can simply be done by global thresholding with the  $H$  and  $V$  channel of the HSV color model. Small regions not belonging to the ruler can also remain by this operation, so we take only the largest region in the computed mask (Figure 2.1(a) and Figure 2.1(b)) .

**2. Rotation into horizontal position:** In the next step we determine the orientation of the region describing the ruler. For this purpose we calculate the angle between the x-axis and the major axis of the ellipse that has the same second-moments as the region (the ellipse with the same orientation as the ruler). Once we have this angle, we can rotate the ruler into a horizontal position (Figure 2.1(c)).

**3. Scanlining:** For robustness we use three scanlines to determine the number of pixels between two marks (a scanline is a vector containing the pixel values of a specific line (y-value) of an image from left to right). The first scanline is on the midpoint between the top and bottom y-value of the ruler. The two others are 10 pixels above and below, respectively (Figure 2.1(d)). To determine the scale we take the maximum number of pixels between two marks in all of the three scanlines. With this method we use the part of the ruler which lies most normal to the camera and has the smallest curvature and therefore has to be the most accurate measurement. Occasionally it can happen that one or more marks are not recognized in the scanline (e.g. when the ruler has a strong curvature), hence too large distances with more than 200 pixels between two marks are rejected (in all images used by us no greater distance than 155 pixels could be found). If  $n$  denotes the number of pixels corresponding to one centimeter, the area of one pixel in the image is  $100/n^2 \text{ mm}^2$ . The example given in Figure 2.1 results in a maximum distance of 138 pixels between two marks, i.e. one pixel of the image has an area of  $\sim 0.00525 \text{ mm}^2$ .

## 2.2 Segmentation of Hemangiomas

The task of the segmentation algorithm is to find the regions in an image belonging to the hemangioma. In Section 2.2.1 we give an overview of several existing methods proposed in the literature dealing with the segmentation of skin lesions, explain the shortcomings of them for our purpose and finally justify the use of our classification-based segmentation. After that our segmentation method is described in detail, divided into a preprocessing step (Section 2.2.2) and a classification step (Section 2.2.3).

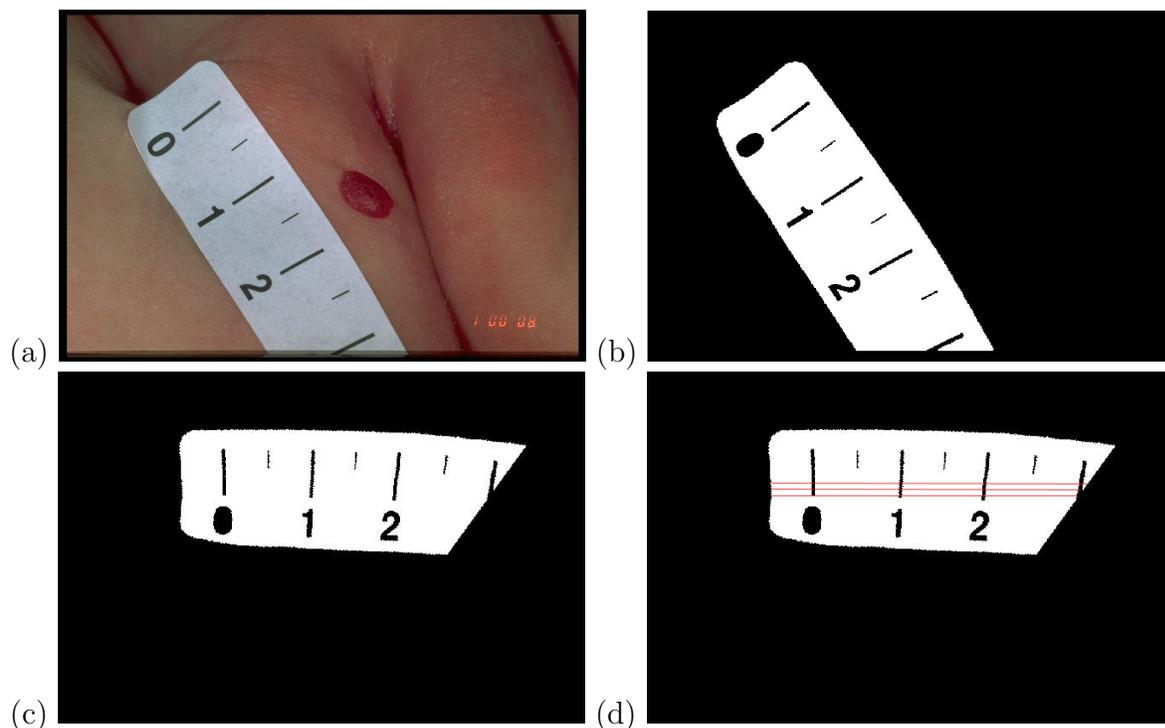


Figure 2.1: Example for determining the scale of an image.

### 2.2.1 Evaluation of Methods for Skin Lesion Segmentation

**Global thresholding:** A thresholding operation is used for the segmentation of skin lesions in [13] and [16]. However, finding reliable parameter values for thresholding operations without user interaction working well with all of our images cannot be done. This is mainly caused by the low contrast between skin and hemangioma regions, since hemangiomas mainly appear in red.

**Intensity image thresholding:** A more promising method for our purpose is given in [41]. Here an image showing a melanoma is first transformed into an intensity image where the intensity at a pixel shows the visual difference of that pixel to the background (i.e. the skin). Then a threshold value is determined from the average intensity of the  $p\%$  highest gradient pixels in the intensity image obtained to find approximate lesion boundaries. Finally the lesion boundary is refined using edge information in the image. The problem of this algorithm for our task is choosing an eligible value for  $p$ . For an image containing a large hemangioma this value should be taken larger than for an image containing a small lesion. For a set of images with small variance of lesion sizes a failure made here can be compensated in the refining step. But hemangioma sizes vary strongly from one image to another. In Figure 2.2 this problem is demonstrated with an image containing a small spot-shaped hemangioma (Figure 2.2(a)) and an image containing a comparatively large hemangioma (Figure 2.2(b)). For the small hemangioma image a percentile  $p=0.01\%$  looks adequate (Figure 2.2(c)). Segmenting the other image with

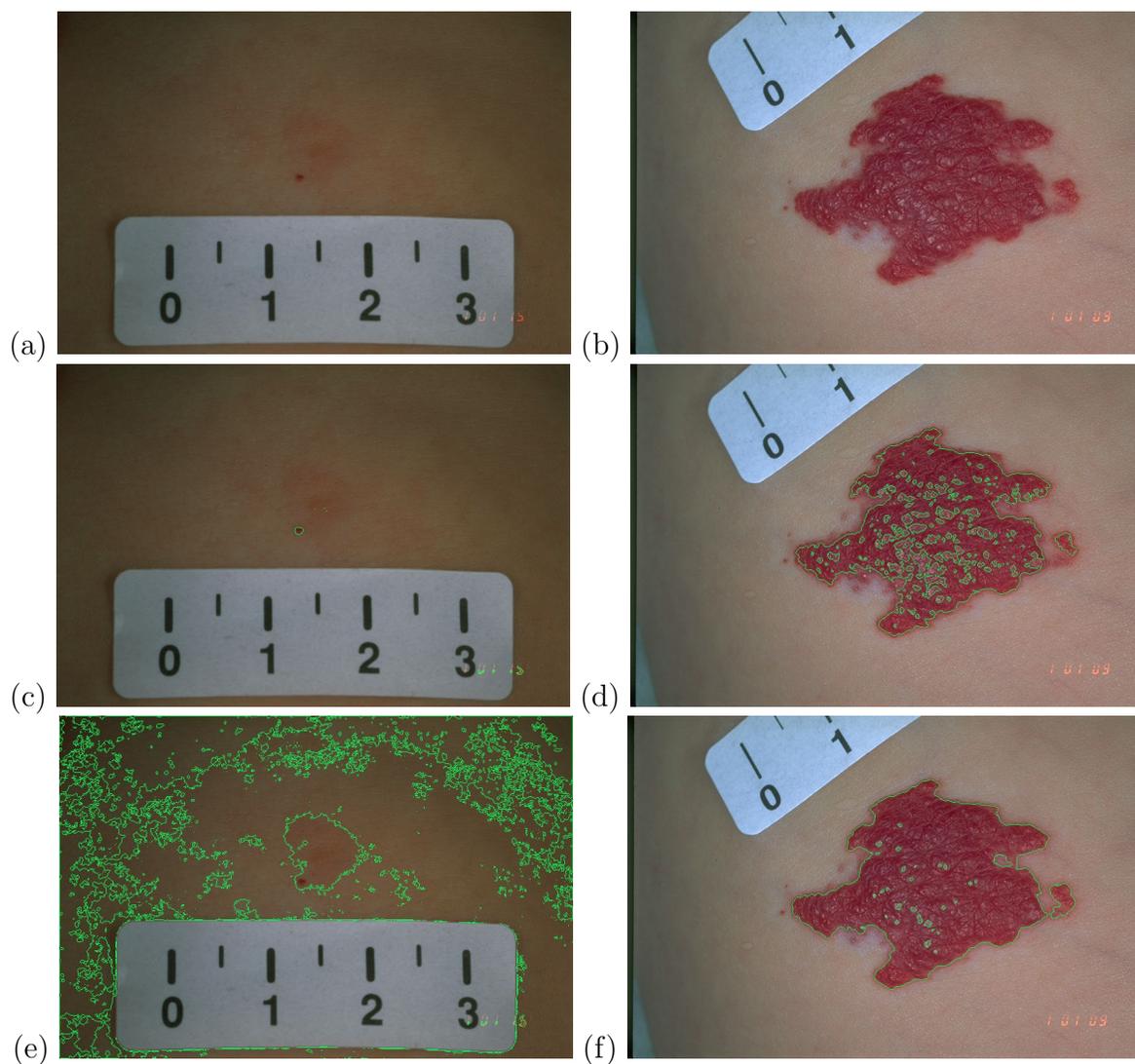


Figure 2.2: Results of intensity image thresholding of two images with different values for  $p$ . The results in (c)/(d) and (e)/(f) are obtained with the same parameters, respectively. Note, each of the two parameter values gives satisfying results only for one of the images.

this percentile detects the borders too far inside the hemangioma (Figure 2.2(d)). For that image a percentile  $p=25\%$  is optimal (Figure 2.2(e)). However, this value is too large for the image containing the small hemangioma (Figure 2.2(f)). Nevertheless, the intensity image describing the visual difference between skin and hemangioma is a useful feature for our purpose and is used slightly modified as a feature for the classification step.

**Region-based methods:** Further possibilities for segmenting skin lesions are region-based segmentation algorithms. With region-based algorithms not only the color values of pixels but also their spatial relationships are considered. Split-and-merge is a widely used region-based segmentation algorithm [21]. It first splits an image into homogeneous

sub-regions and then merges together regions with similar average color values. Round et al's work on segmentation of skin lesions is basically an application of the split-and-merge algorithm [36]. Schmid and Fischer presented a region-based approach working with two-dimensional histogram analysis and the fuzzy c-means clustering technique [38]. For our purpose we tested the region-based method explained in [17] on our images. The algorithm proposed in this work first splits the image into smaller regions until all the regions meet the homogeneity criteria set by a threshold. Second, the small split regions are grouped by DBSCAN clustering algorithm to form the final regions. The results of applying this method on two of our images are shown in Figure 2.3.

The method was rejected due to the reason that after the clustering process it is very difficult to decide which clusters belong to the hemangioma and which do not. Although the clustering of Figure 2.3(a) induces a large cluster containing the major part of the hemangioma (Figure 2.3(b), note that the black regions are detected to be non-skin before, see Section 2.2.2), there are several clusters near the border of the hemangioma which cannot be classified easily. A solution to this problem would be to change the parameters of the algorithm to merge together regions having more different color values. This would give us a single cluster containing the whole hemangioma. But although the image of Figure 2.3(c) was clustered with the same parameters as the image of Figure 2.3(a), the result (Figure 2.3(d)) shows clusters where skin and parts of the hemangioma were merged together. It was impossible to find a set of parameters working well with most of the images.

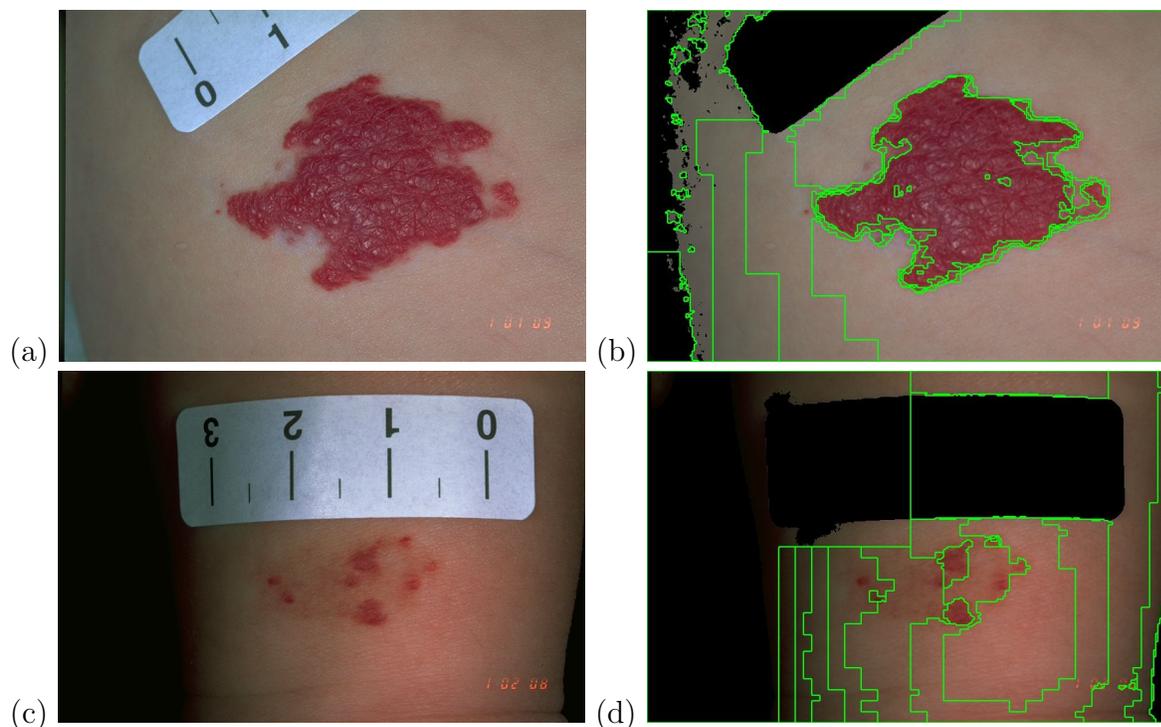


Figure 2.3: Result of DBSCAN clustering of two images.

As a conclusion, the problem of all color segmentation algorithms is to determine

which of the segmented regions are hemangioma regions. Making use of specific a priori knowledge about the general appearance of hemangiomas is difficult because of several reasons:

- Hemangiomas can consist of several unconnected regions, so more than one region has to be classified as hemangioma.
- Hemangiomas do not have a specific shape.
- The hemangioma size in the images differs to a high extent.

Due to the difficulties faced by existing approaches mentioned previously we propose a different method: we use a classifier for segmenting the images that classifies each pixel in the image as hemangioma or non hemangioma on the basis of its color values and the surrounding skin region. In the following we describe the process to determine the type of classifier and set of features suitable for our purpose.

## 2.2.2 Preprocessing

Before the classification step the images have to be preprocessed to improve the accuracy of the classifier and to reduce computation time. To remove noise a median filter is applied on the images. Furthermore, image regions containing no skin are masked out and the images are normalized in such a way that skin has nearly the same color values in all images, i.e. the skin serves a reference color to account for different illumination situations during image acquisition.

### Noise Removal

The task of the noise removal process is to reduce small structures in the images. Smoothing the images causes less color variations in the skin and hemangioma and makes thereby the segmentation process more robust. With low-pass filtering an image is smoothed by replacing each pixel by a weighted sum of its neighbors. However, with median filtering, the value of an output pixel is determined by the median of the neighborhood pixels. The advantage of this method in contrast to low-pass filtering is that it preserves the edges of the image while reducing the noise. In this work we use a median filter with a window size of  $5 \times 5$ , i.e. the value of an output pixel is determined by the median of all pixel included in  $5 \times 5$  window surrounding the particular pixel. By testing this window size turned out to be the best compromise between reducing noise and preserving relevant information in the images (note that the window size of  $5 \times 5$  is significantly smaller than the typical hemangioma size of appr.  $200 \times 200$  pixels).

### Non-Skin Masking

A simple test for masking out non-skin regions is used after the median filtering to exclude regions that likely are not part of the skin or the hemangioma (e.g. the ruler or dark areas in the background). This step is necessary for a robust determination of the skin color. Our method is based on a heuristic proposed in [19] but substantially simpler. We only

check two criteria for each pixel ( $R$ ,  $G$  and  $B$  are the red, green and blue color values of the RGB color model):

1.  $R < G$
2.  $R < B$

If one of the criteria is satisfied, the particular pixel is marked as non skin. This test makes use of the fact that skin has usually a reddish color and therefore shows a greater red portion than green and blue portion. By applying this method on 122 test images we got an almost perfect result for 96 images (classification error less than 1%). The rest shows an average classification error of about 5%. The non-skin masking never causes problematic results for the subsequent steps, because the hemangioma and the bigger part of the skin are never masked out. Figure 2.4 shows the non skin masking results of three particular images (non skin regions are marked blue). Non skin masking of the image in Figure 2.4(a) produces a nearly perfect result (Figure 2.4(b)). Worse results are achieved in Figure 2.4(c) and (e). The failure of Figure 2.4(d) is mainly caused by a red cloth in the dark background of the image. In Figure 2.4(e) some parts of the skin are erroneously masked out in a rather badly illuminated region of the image. Here the skin loses its reddish appearance. Note that these errors do not affect the hemangioma measurement.

### Normalization with Skin Color

To achieve more accurate classification results despite differences of illumination between the images a normalization with the skin color has to be done on the images. We determine the color value of the skin in an image and subtract this value from all pixels with the aim of having nearly the same color value of (0,0,0) for skin pixels in all images. The necessity of this step is shown in Table 2.1 where we have computed the classification result (false positive and false negative rate, for a definition see Section 5.1.3) of a single-layer perceptron (see Section 2.2.3) for 15 images without normalization and compared it with the results achieved with normalization. The same feature set as for the classification step (see Section 2.2.3) was used and all images were reduced to a resolution of 256x384 and randomly divided into a training set (30%) and a test set (70%). For normalization we have tested three different variants of computing the skin color:

1. **Manual Normalization:** the user chooses three 3x3 windows near the border of the hemangioma and the mean color value of all windows is chosen to be the skin color (this is only for testing, since we want an automated procedure with no user interaction).
2. **Median Normalization:** the median value for each color channel in the image is chosen.
3. **Histogram Normalization:** a 3D histogram of the RGB color channels is created and the maximum RGB value, which has a brightness greater than 120, is chosen. This threshold was decided to be the most appropriate by empirical tests.

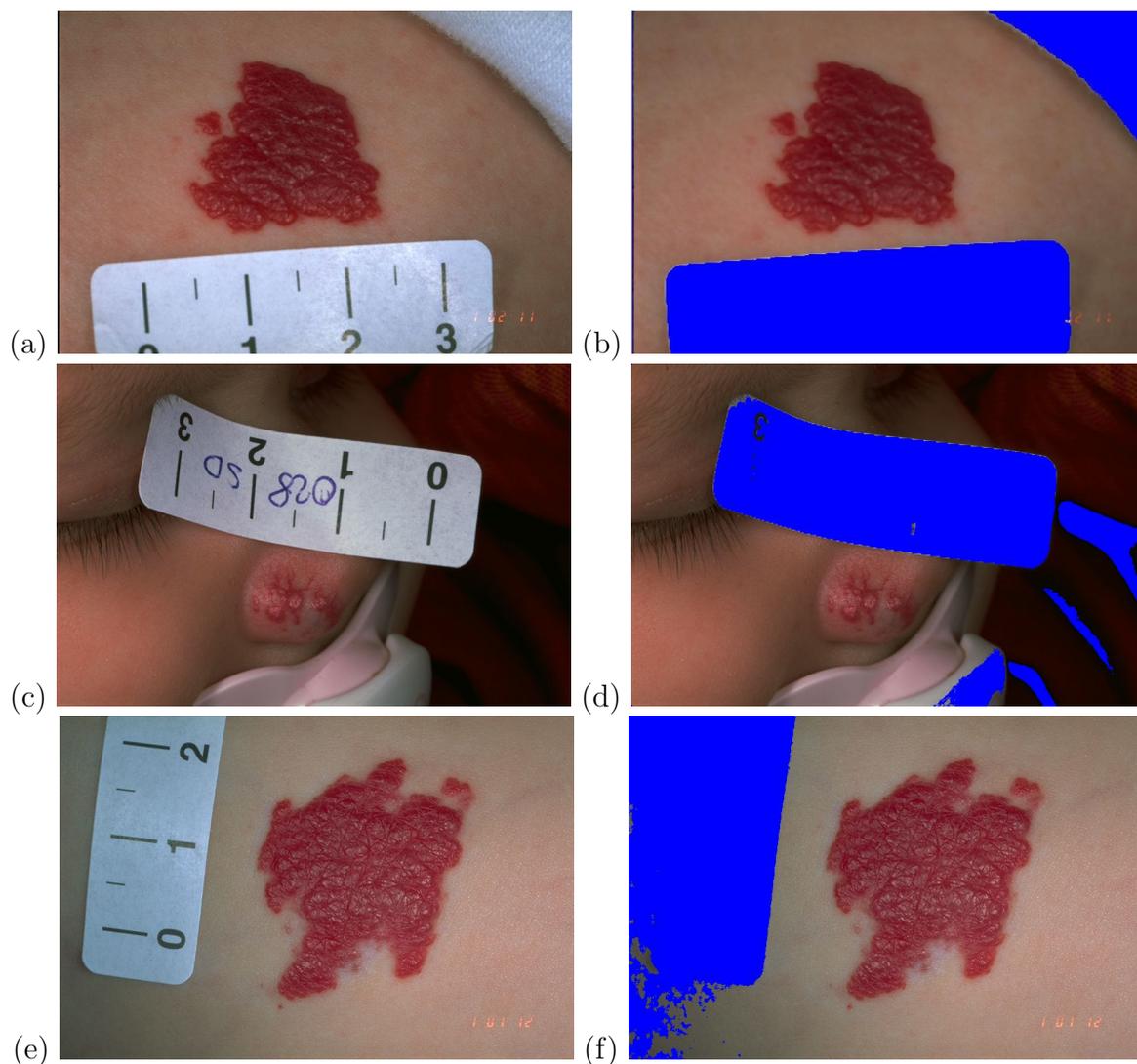


Figure 2.4: Non skin masking results of three images.

Both Median and Histogram Normalization make use of the fact that after non skin masking the majority of the remaining pixels in an image represents skin. As can be seen in Table 2.1, without normalization the perceptron is not able to find a useful decision boundary. The perceptron has set a decision boundary that classifies nearly all pixels as non hemangioma and therefore we get a false positive rate of 0.0% and a false negative rate of 99.8% without normalization. With normalization we obtain more usable decision boundaries resulting in a total error rate (false positive rate plus false negative rate) of 20.8% (Manual Normalization), 21.6% (Median Normalization) and 20.5% (Histogram Normalization) in the test set. Although median and histogram normalization accomplish nearly the same results, the subjective visual result on particular images is better with histogram normalization. It can also be seen that the manual normalization does not achieve better results than the histogram normalization (20.5% against 20.8% total error rate in the test set) which is an indication that the histogram normalization works

	Test Set			Training Set		
	False Pos.	False Neg.	Sum	False Pos.	False Neg.	Sum
No Norm.	0.0%	99.8%	<b>99.8%</b>	0.0%	99.8%	<b>99.8%</b>
Manual Norm.	6.0%	14.8%	<b>20.8%</b>	6.0%	14.5%	<b>20.5%</b>
Median Norm.	6.1%	15.5%	<b>21.6%</b>	6.0%	15.2%	<b>21.2%</b>
Histogram Norm.	5.8%	14.7%	<b>20.5%</b>	5.8%	14.4%	<b>20.2%</b>

Table 2.1: Classification results after various normalization methods.

accurately.

### 2.2.3 Pixel-wise Classification

As mentioned before, segmentation is achieved by a pixel-wise classification based on the pixel’s color values. In this section first various classifiers are discussed and the final choice for the single-layer perceptron is argued. Next, a detailed description of its classification and training technique is given. Finally, the selection of an appropriate set of features is described.

#### Choice of the Classifier

To get satisfactory classification results the data to be classified (i.e. the color values) have to show some form of clusters, where a cluster describes data values belonging to the same class. To test that we first normalized the RGB color values of 10 images in such a way that the skin has nearly the same color values in all images (see Section 2.2.2). By plotting their RGB values in 3D space (Figure 2.5(a), the side views are shown in Figure 2.5(b)-(d)) it can be seen that the color values of the hemangioma (red) are mostly separated from the other values (blue).

**Single-layer perceptron:** The single-layer perceptron is a simple type of binary classifier having the advantage of a simple and fast classification [35]. Looking at it geometrically a single-layer perceptron classifies the data of Figure 2.5 by putting a plane in the 3D space that divides the data set into two classes. Therefore an adequate basic segmentation result can be assumed on our images by applying a single-layer perceptron. As we see later, this assumption is confirmed, especially if we take more appropriate color values from other color models.

**Neural network:** Beside the single-layer perceptron various other classifiers exist. A neural network (multi-layer perceptron) is an improvement of the traditional single-layer perceptron with the ability of setting non-linear decision boundaries instead of only linear ones [2]. In other words, a neural network can put a curved surface in 3D space as decision boundary. We have tested neural networks for segmenting our images with the conclusion that the results are not better than by segmenting with a single-layer perceptron. Since a single-layer perceptron is simpler and faster in classification, we kept the perceptron.

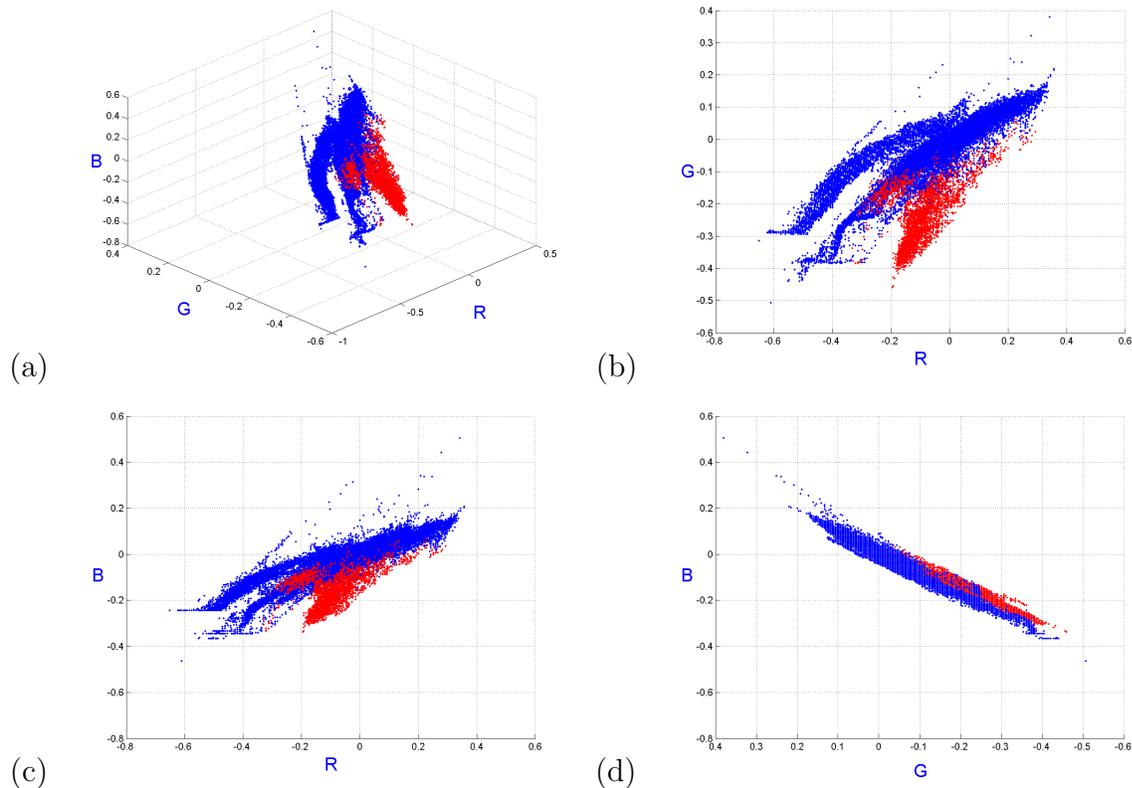


Figure 2.5: 3D plot of normalized RGB color values of 10 images. The red points represent the color values of the hemangioma pixels.

***k*-Nearest Neighbor classifier:** The *k*-Nearest Neighbor (*k*-NN) classifier determines the *k* nearest feature vectors of a reference set for every feature vector [7]. The specific feature value is mapped to the class having the majority in the *k* nearest feature vectors. Due to the complex feature distribution a *k*-NN classifier with *k* = 5 achieved better segmentation results than the single-layer perceptron on our images. The *k*-NN classifier achieved an average false positive rate of 1.8% and an average false negative rate of 16.1% on 7 images, where the perceptron achieved an average false positive rate of 5.3% and an average false negative rate of 16.9% on the same 7 images. Nevertheless, the *k*-NN classifier is not practical for our purpose because of its high computation time, since every pixel of the image has to be classified individually.

Due to the similar results of the neural network and the bad computation time of the *k*-NN classifier we decided to use the single-layer perceptron for classification.

### The Single-Layer Perceptron

The single-layer perceptron is a simple binary classifier based on a mathematical model for the behavior of a single biological neuron.

Assume that we have a *d*-dimensional feature vector  $x \in \mathbb{R}^d$  and two classes  $\omega_1, \omega_2$ . Our goal is to find a mapping  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  with

$$g(x) > 0 \text{ if } x \in \omega_1 \quad (2.1)$$

$$g(x) < 0 \text{ if } x \in \omega_2 \quad (2.2)$$

$g$  is called the *discriminant function* and has the form

$$g(x) = \sum_{i=1}^d w_i x_i - \theta = \mathbf{w}^T \mathbf{x} - \theta \quad (2.3)$$

where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_d \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_d \end{pmatrix} \quad (2.4)$$

$\mathbf{w}$  is called the weight vector and  $\theta$  the *bias*. If we call  $o(\mathbf{x})$  the output of the perceptron, we have

$$o(\mathbf{x}) = g(\mathbf{w}^T \mathbf{x} - \theta) = \begin{cases} 1 & \text{if } \mathbf{w}^T \mathbf{x} \geq \theta \\ -1 & \text{if } \mathbf{w}^T \mathbf{x} < \theta \end{cases} \quad (2.5)$$

where  $g()$  is the signum function:

$$\text{sgn}(a) = \begin{cases} +1 & \text{if } a \geq 0 \\ -1 & \text{if } a < 0 \end{cases} \quad (2.6)$$

It is common practice to “pull” the bias into the weight vector by introducing additional coordinates  $x_0 = 1$  and  $w_0 = -\theta$ .

$${}^a \mathbf{x} = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \dots \\ x_d \end{pmatrix}, {}^a \mathbf{w} = \begin{pmatrix} -\theta \\ w_1 \\ w_2 \\ \dots \\ w_d \end{pmatrix} \quad (2.7)$$

## Perceptron training

Let  $S_{Tr} = \{\mathbf{X}, \mathbf{t}\}$  denote a set of  $N$  augmented input vectors  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N) \in \mathbb{R}^{(d+1) \times N}$  and corresponding class labels  $\mathbf{t} = (t_1, \dots, t_N), t_i \in \{1, -1\}$ , called *training set*.

Goal: find an augmented weight vector  $\mathbf{w}$  such that

$$o(\mathbf{x}_i) = \text{sgn}(\mathbf{w}^T \mathbf{x}_i) = t_i, 1 \leq i \leq N \quad (2.8)$$

**Perceptron Learning Rule:** if a training vector  $\mathbf{x}_j$  with  $t_j = 1$  is misclassified, add a multiple of  $\mathbf{x}_j t_j$  to  $\mathbf{w}$ :  $\mathbf{w}_{\text{new}} = \mathbf{w} + \gamma \mathbf{x}_j t_j$ . Likewise, if a training vector  $\mathbf{x}_j$  with  $t_j = -1$  is misclassified, subtract a multiple of  $\mathbf{x}_j t_j$  from  $\mathbf{w}$ :  $\mathbf{w}_{\text{new}} = \mathbf{w} + \gamma \mathbf{x}_j t_j$ . The factor  $\gamma$  is called *learning rate*.

## Feature Selection

For classification we have to define a set of features showing a big difference between skin and hemangioma pixels. Possible features for the classification are at first the color channels of three significant color spaces [23]:

1. **RGB**: In RGB color space, each color is represented by a three number triple. The components of this triple specify, respectively, the amount of red, green and blue in the color.
2. **HSV**: In the HSV color space, each color is again represented by a three number triple. The first component, *Hue*, describes the basic color in terms of its angular position on a “color wheel”. The *Hue* is described in terms of degrees. The second component of the HSV triple is *Saturation*, which can be thought of as the “purity” of the color. The third component is *Value*, which is a measure how “bright” the color is.
3. **CIE 1976 L\*a\*b\***: The CIE 1976 L\*a\*b\* color space was created by the *Commission Internationale de l’Eclairage* and has the special ability that the color differences perceived by the human eye correspond to colorimetrically measured distances. The  $L^*$ -axis in this three-dimensional color space is known as the lightness, the other two coordinates  $a^*$  and  $b^*$  represent redness-greenness and yellowness-blueness, respectively.

Further we have created a 10th feature called *abdist* representing the color distance between healthy skin and the hemangioma (a detailed description is given in the next paragraph). To find an appropriate set of features we have ascertained the false negative and false positive rate for every meaningful group of features on a test set of 15 normalized images (Table 2.2). The images were reduced to a resolution of 256x384 before and randomly divided into a training set (30%) and test set (70%).

It can be seen that the set  $\{G, H, a^*, abdist\}$  achieves the best results on the test set and therefore these features are used for classification. Their use is also justified in Figure 2.6 where the intensity images of the four selected features are shown. It is apparent that each feature has an observable difference between pixels belonging to the hemangioma and pixels belonging to the skin.

It must be noticed that for classification the  $H$ -values have to be rotated 180 degrees to facilitate computation with the circular data. In the HSV color model the red color value lies at  $0^\circ$ , i.e. similar red color values can have very different  $H$ -values. Since hemangiomas and skin have  $H$ -values in this area, they have to be transformed into a linear range. Although many sophisticated methods for handling circular data are covered in literature (a detailed overview is given in [12]), for our task a simple rotation of 180 degrees is effective, because the red area is the only relevant one for the classifier.

**The Feature *abdist***: The feature *abdist* stands for the Euclidean distance between the skin and the hemangioma in the  $L^*a^*b^*$  color space without consideration of the luminance  $L^*$  and intensification of the  $a^*$  component. This feature is adopted from

Features	Test Set		Training Set	
	False Pos.	False Neg.	False Pos.	False Neg.
$L a^* b^*$	4.6 %	37.9 %	4.6 %	37.2 %
$abdist$	6.8 %	25.2 %	6.8 %	25.0 %
$a^*$	6.0 %	31.7 %	6.0 %	31.3 %
$G H a^*$	6.0 %	34.9 %	6.0 %	34.3 %
$H S a^*$	4.7 %	46.3 %	4.7 %	45.5 %
$V abdist$	5.6 %	27.0 %	5.5 %	26.6 %
$G H a^* abdist$	4.4 %	14.7 %	4.4 %	14.4 %
$R G H S V a^*$	5.0 %	41.5 %	5.0 %	41.0 %
$H S abdist$	4.3 %	23.9 %	4.3 %	23.6 %
$all$	4.4 %	17.9 %	4.4 %	17.6 %

Table 2.2: False positive and false negative rate of a set of feature groups of 15 images trained with a perceptron.

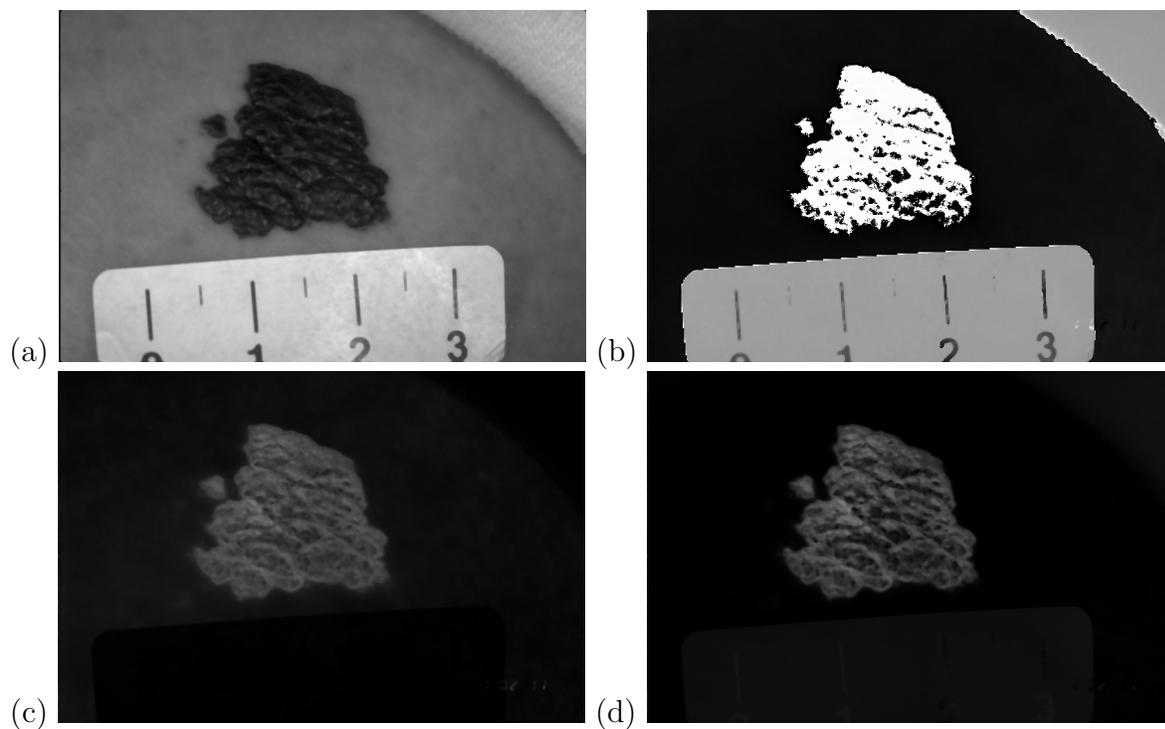


Figure 2.6: The intensity images of the four features (a)  $G$ , (b)  $H$ , (c)  $a$  and (d)  $abdist$  of the image in Figure 2.4(a).

[41] (see Section 2.2.1). In this paper the proposed method works on an intensity image describing the Euclidean distance between the skin and the lesion.

The advantage of the *CIE 1976 L\*a\*b\** color space compared with other color spaces is that it defines color in such a way that the Euclidean distance between two colors is proportional to their visual difference. This property is very useful for our purpose, since the color distance in RGB space between hemangioma and skin regions is not as high as the perceived difference.

If  $a_s, b_s$  denotes the  $a^*$  and  $b^*$  values of the skin (obtained from the normalization step, see 2.2.2) and  $a_p, b_p$  that from a particular pixel, its *abdist* is computed as follows:

$$abdist = \sqrt{(2a_s - 2a_p)^2 + (b_s - b_p)^2} \quad (2.9)$$

The difference of the  $a^*$  channel is multiplied by the factor 2, because the  $a^*$  value differs more between hemangioma and skin pixels than the  $b^*$  value. Finally for contrast enhancement we apply a Gaussian function of the form  $G(x) = 1/\sqrt{2\pi\sigma}(1 - \exp(-x^2/2\sigma^2))$  with  $\sigma = 0.5$  on the intensity image. As can be seen in Figure 2.7, this function increases higher values (hemangioma) and decreases lower values (skin) in the intensity image. For an example of *abdist* see Figure 2.6(d).

### Treatment of Highlights

Highlights on the hemangioma are normally erroneously detected as healthy skin by the classifier. This is corrected by closing all holes occurring in the masked region. Since hemangiomas with large holes of normal skin could not be found in the data, they seem to be very rare and a possible error resulting from this operation can be neglected. The effect of hole closing can be seen in Figure 2.8: the highlights detected as non hemangioma by the classifier (Figure 2.8(a)) are added to the mask by hole closing (Figure 2.8(b)).

In Figure 2.9 two additional final segmentation results of images taken at the first examination during follow-up studies are shown. The detected hemangioma region is

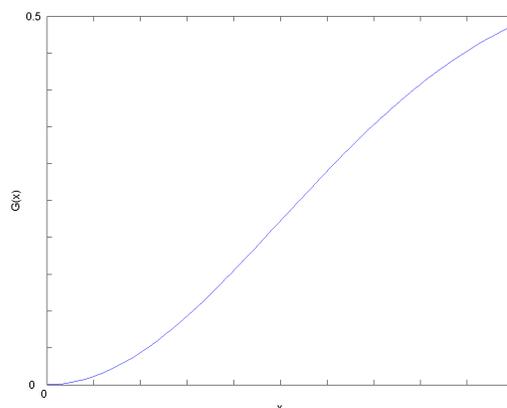


Figure 2.7: Gaussian function  $G(x) = 1/\sqrt{2\pi\sigma}(1 - \exp(-x^2/2\sigma^2))$  with  $\sigma = 0.5$

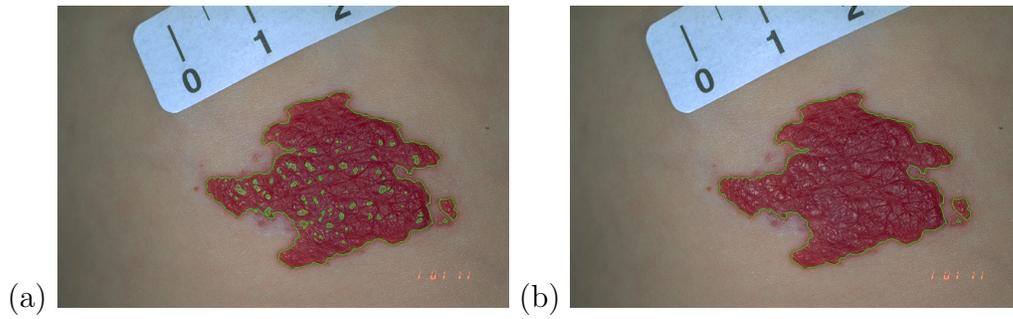


Figure 2.8: Result of hole closing on a particular image of a hemangioma.

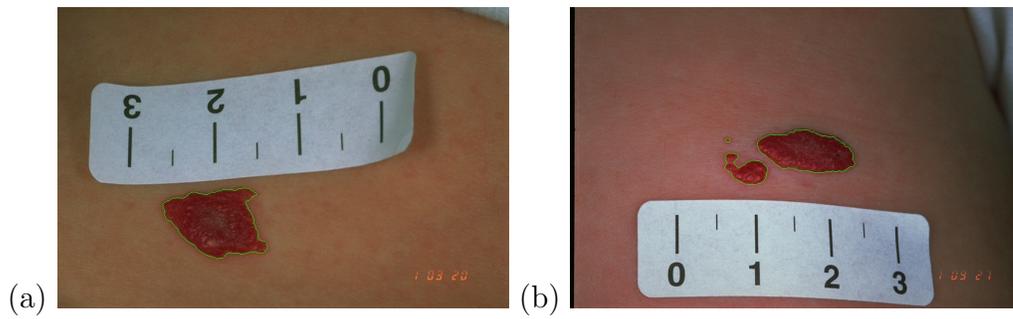


Figure 2.9: Segmentation result on images taken at the first examination during follow-up.

used to measure the area of the affected skin at the start of treatment. Moreover, using the image registration method presented in Chapter 3, this region is aligned to the new hemangioma region in subsequent images of follow-up. The transformed region allows for a quantification of size decrease or increase and, additionally, serves as a region of interest for the regression detection.

# Chapter 3

## Registration of Follow-Up Hemangioma Images

This chapter presents an automatic method for registering hemangioma images taken during clinical trials at specific time intervals. The registration of follow-up images is a necessary prerequisite for the detection of regressions (see Chapter 4). It allows for the establishment of correspondences across the image series, and thereby enables the propagation of the initial hemangioma shape to subsequent images. This makes a definition of a region of interest for the regression detection possible, allows for shape and area comparisons, and finally for a direct comparison of corresponding color values. This is essential for the detection of regression regions exhibiting subtle color differences.

The chapter starts with a general introduction to the field of image registration in Section 3.1. In Section 3.2 an overview of our method is given. The method is presented in detail in the following sections, divided into the detection and description of interest points (Section 3.2.1), the matching of interest points (Section 3.2.2) and the estimation of the transformation between the two images (Section 3.2.3).

### 3.1 Introduction to Image Registration

Image registration is the process of geometrically aligning two images taken from different viewpoints [45] and, in our case, also taken at different times. In the image registration process one image is kept unchanged (the *reference image*) and the other one (the *sensed image*) is transformed to be aligned with the reference image. Image registration methods can be characterized in various ways [45]: concerning the transformation model, a registration can either be *rigid* or *non-rigid*. With rigid registration the transformation between the two images is globally defined by a linear transformation, whereas non-rigid registration allows local deformations. Additionally, image registration can be specified by the type of images or acquisition process: images to be registered can originate from different viewpoints (*multiview analysis*), different times (*multitemporal analysis*) and different sensors (*multimodal analysis*). In terms of the basic methodology, registration algorithms are divided into *feature-based methods* and *area-based methods*. Both principles are explained in the next sections.

### 3.1.1 Feature-Based Methods

Feature-based methods use distinctive interest points apparent in both images for estimating the transformation from the sensed to the reference image. The registration process is shown with a simple example in Figure 3.1. The following four steps are executed:

1. **Feature detection:** Interest points are detected in both images. In Figure 3.1 the smiley's eyes and corners of the mouth are detected as interest points in both images (red crosses).
2. **Feature matching:** Interest points are matched by means of their feature descriptions (blue lines).
3. **Transform model estimation:** Matched interest points are used to compute the parameters of the mapping function. In the example of Figure 3.1 the four correspondences can be taken to estimate the scaling and translation of the smiley from the sensed to the reference image.

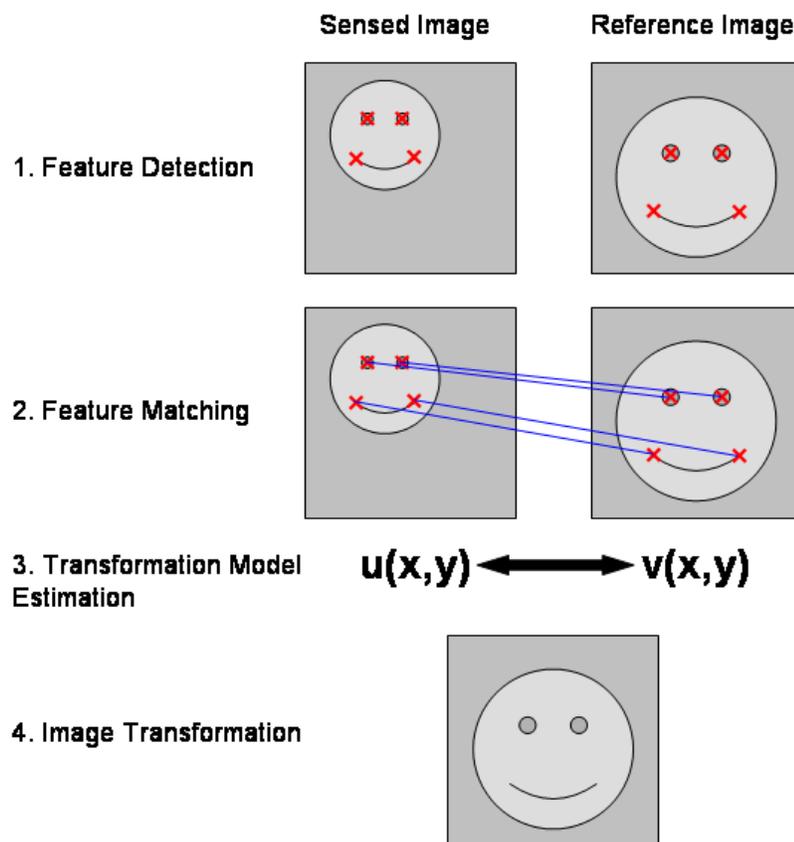


Figure 3.1: The four steps of feature-based image registration methods.

4. **Image transformation:** The image is finally transformed using the computed mapping function and usually bilinearly interpolated.

### 3.1.2 Area-Based Methods

With area-based registration methods no image features are initially detected. Their principle is to combine the feature detection and feature matching step by comparing small windows around points in the reference and sensed image. Correspondence of image points is based on the similarity measure between two given windows, e.g. the normalized cross-correlation, correlation coefficient or sequential similarity detection [45]. Although widely used, area-based methods are not well-adapted for the registration of changing structures like hemangiomas since similarities of corresponding image points become low in changing image regions of follow-up images.

### 3.1.3 Methods used for Skin Lesion Registration

Several approaches for the registration of dermatological images were proposed in the past dealing mainly with images of skin lesions representing melanoma or psoriasis. Maglogianis [26] and Pavlopoulos [33] both propose a similar hybrid method using the log-polar transformation for estimating scaling and rotation parameters and a sign change similarity criterion in combination with a hill-climbing optimization scheme for translation estimation. Ersbøll et al. [30] work with statistical shape analysis after lesion segmentation to do a first rigid alignment under the assumption of image scale constancy. Afterwards small internal displacements are corrected by a combined registration and alignment scheme. The SHARP-algorithm presented in [9] segments the lesions in the images and uses the first and second order moments of the resulting binary images to determine the rotation and translation parameters.

## 3.2 Robust Feature-Based Method for Registration of Follow-Up Hemangioma Images

The methods outlined in Section 3.1.3 are not capable of handling certain properties of typical hemangioma follow up images: differences of view point during image acquisition typically exceed a range that can be accounted for by 2D rotations and translation in the image plane, the change in image scale can be high, and hemangiomas change appearance during the period of follow-up examinations, making a robust approach necessary. In Figure 3.2 these cases are illustrated: Figure 3.2(a) and Figure 3.2(b) show a pair for which reliable matching is only possible for a small part of the hemangioma due to the healing process between examinations, and rotation between images of appr.  $170^\circ$ . Figure 3.2(c) and Figure 3.2(d) are examples for a large scale change, and Figure 3.2(e) and Figure 3.2(f) show a pair of images where the viewpoint change is high.

Because of these issues, our approach is based on the detection and matching of distinctive interest points by means of local features. The resulting matchings of interest points are

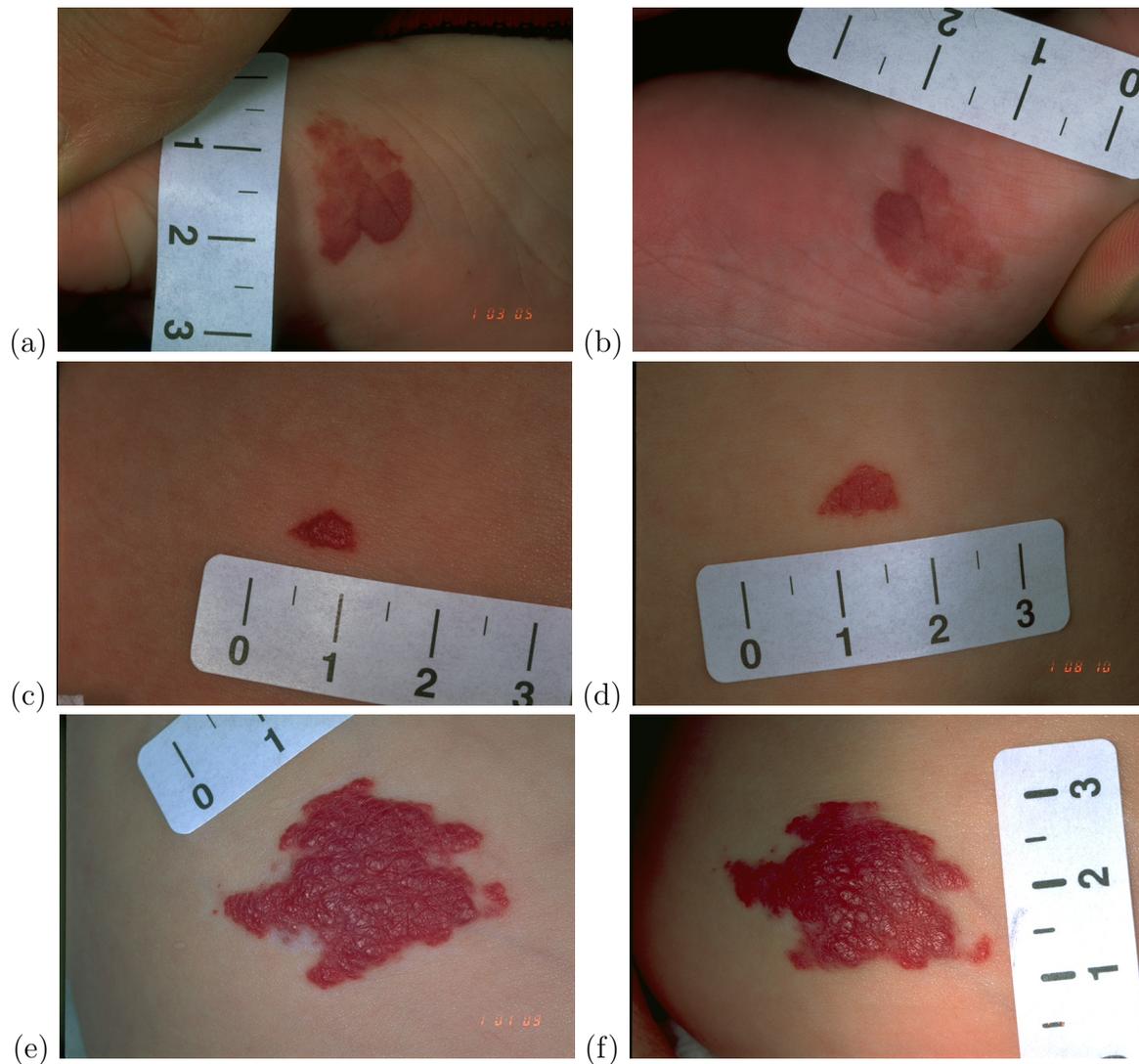


Figure 3.2: Three image pairs representing images of the same hemangioma taken at different times.

used for a robust estimation of the transformation between images by RANSAC. To obtain reliable matches of interest points we use SIFT features [25] for description of interest points and determination of correspondences between images. Under the assumption that hemangiomas are planar, the transformation between images is defined as a homography which is estimated by the detected point correspondences. The basic methodology is illustrated in Figure 3.3.

The method has several advantages over existing approaches for dealing with hemangioma data: SIFT is a rotation and scale-invariant descriptor which allows the reliable matching of points under different views, the modeling of the transformation between images with a homography can represent any projective relation between the images, and finally RANSAC can cope with partially incorrect or missing matches between interest points caused by changed hemangioma appearances.

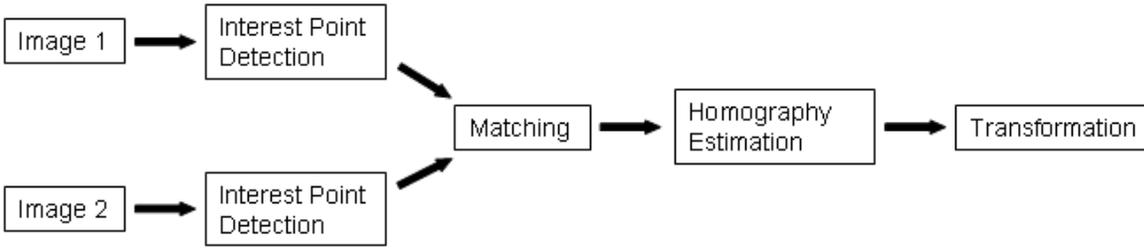


Figure 3.3: Illustration of the feature-based method used for hemangioma image registration.

### 3.2.1 Detection of Interest Points in Hemangioma Images

As mentioned above, the first task is to find distinctive interest points in the images to be registered. In this section first the original SIFT method and the Canny edge detector are described. Afterwards, the use and adaptation of these methods for interest point detection in hemangioma images is presented and reasoned.

#### Scale Invariant Feature Transform

The scale-invariant feature transform (SIFT) presented by Lowe [25] is a scale and rotation invariant local descriptor of features in images and has proven to be a robust and therefore widely used method [31]. It consists of four major steps:

1. **Scale-space extrema detection:** The detection of interest points works on different octaves, i.e. different resolutions of the image. For each octave a Gaussian scale space is computed by repeatedly convolving the image with a Gaussian filter kernel and subsequently Difference-of-Gaussians images are created by subtracting the blurred images with adjacent scales. Candidate locations for interest points are then detected at extrema of the Difference-of-Gaussians images.
2. **Keypoint localization:** For each initial keypoint (interest point) position an interpolated more exact position is determined by a 3D quadratic function. Furthermore, unstable extrema with low contrast are rejected.
3. **Orientation assignment:** In order to achieve rotation invariance of interest point descriptors, to each interest point a main orientation determined by means of peaks in an orientation histogram of surrounding gradient orientations is assigned. All future operations are performed relative to the main orientation.
4. **Keypoint description:** Again the local gradient data is used to create a set of histograms over a window centered on the keypoint. Usually a set of 16 histograms, aligned in a 4x4 grid, each with 8 orientation bins is used. This results in a feature vector of 128 elements for each interest point.

## Canny Edge Detector

The aim of an edge detector is to provide a binary version of the image where the detected edges represent the object's borders. The Canny edge detector was presented by J. Canny in 1986 and was designed to be an optimal edge detector regarding the tradeoff between the detection and localization of edges [3]. It consists of the following steps:

1. **Noise removal:** Convolution of the image  $I(x, y)$  with a gaussian kernel  $G_\sigma$

$$I_s = G_\sigma * I \quad (3.1)$$

2. **Gradient magnitude computation:** Computation of gradients  $\nabla I_s$  and gradient magnitudes  $m(x, y)$  of the smoothed image  $I_s$

$$\nabla I_s = \left( \frac{\partial I_s}{\partial x}, \frac{\partial I_s}{\partial y} \right) \quad (3.2)$$

$$m(x, y) = \sqrt{\left( \frac{\partial I_s}{\partial x} \right)^2 + \left( \frac{\partial I_s}{\partial y} \right)^2} \quad (3.3)$$

3. **Non-Maximum suppression:** Detection of local maxima of gradient magnitude by comparison with gradient magnitudes along gradient direction. If  $\frac{\nabla I_s}{\|\nabla I_s\|} = (\delta x, \delta y)$  denotes the unit gradient vector, then a pixel at location  $(x, y)$  is marked as local maximum if

$$m(x, y) > m(x + \delta x, y + \delta y) \quad (3.4)$$

and

$$m(x, y) > m(x - \delta x, y - \delta y). \quad (3.5)$$

4. **Hysteresis thresholding:** Thresholding with a high threshold  $T_{hi}$  and a low threshold  $T_{lo}$  to identify the final edge pixels. A local maximum obtained by the non-maximum suppression is preserved as an edge pixel if its gradient magnitude is above  $T_{hi}$  or its gradient magnitude is above  $T_{lo}$  and one of its neighboring pixels is an edge pixel.

## Interest Point Detection in Hemangioma Images

For our task the standard SIFT method has to be modified on some points, to account for the specific nature of the image data: the interest point detection is constrained to a region covering the hemangioma, only the green color component of the images is taken into account, and Canny edges are used for the interest point localization.

**1. Constraining interest points to the hemangioma region:** The analysis is constrained to the segmented hemangioma region, provided by the method described in Section 2.2. Only interest points inside or near the hemangioma region are accepted. This ensures that the planarity assumption for the homography estimation gives a good estimate of the actual surface, accounts for a better chance of reliable matches in the close vicinity of the hemangioma and reduces computational costs.

In detail, interest points are only accepted if:

1. they are located in the frame encompassing the segmented hemangioma region, enlarged by 40 pixels on each side.
2. they are not located in a background region, i.e. not located in a region that has been detected as non skin by the heuristic described in Section 2.2.2. Note that thereby also the ruler is excluded and differences in its placement do not affect the registration.

**2. Local image representation:** The standard SIFT method works on intensity images. In our case the green channel of the RGB-images is the most reasonable choice. As mentioned in Section 2.2.3, this representation shows higher differences between hemangioma regions and healthy skin which can also be seen in Figure 3.4 where the gray-value representation (mean value of all color channels) and the green channel representation of a hemangioma image is shown. Therefore, the whole registration process works on green channel images.

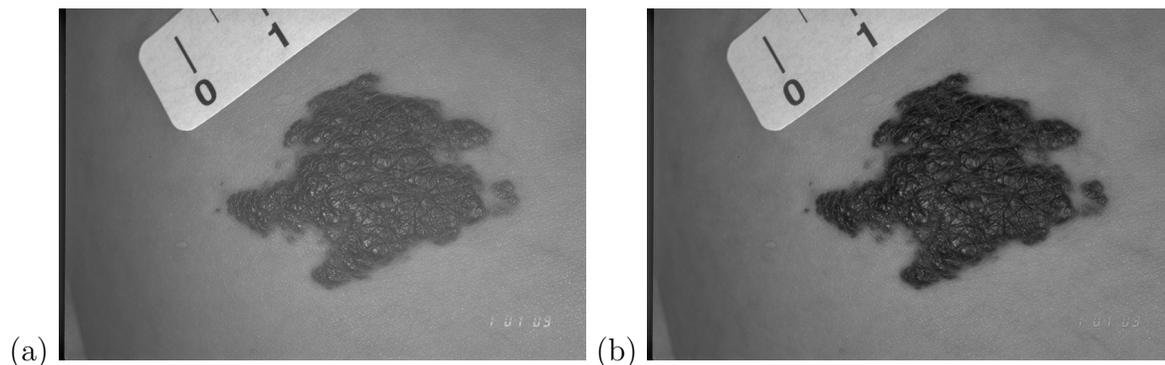


Figure 3.4: (a) Mean value of all color channels and (b) green channel representation of a RGB hemangioma image.

**3. Interest point localization:** In the original SIFT method interest points (called *keypoints*) are localized at extrema in the Difference-of-Gaussians scale space. However, this interest point localization procedure generates an unsatisfactory number of distinctive interest points and correct matches on our images. The reason is that Difference-of-Gaussians tends to have more extrema at intensity variations inside the hemangioma than at the exact hemangioma border. However, the most reliable interest points lie at the

hemangioma border because the inner hemangioma parts change to a higher degree from one time to another. Therefore, in our method interest points are detected along edges in the images. Edge detection is accomplished by means of the Canny edge detector [3] and interest points are finally localized at edge pixels showing the highest gradient magnitude in a neighborhood of  $N$  pixels. By empirical tests a value of  $N = 3$  combined with  $T_{lo} = 0.1$  and  $T_{hi} = 0.2$  (see Section 3.2.1) has proven to be adequate for obtaining a high number of distinctive interest points. The improvement achieved by this method is evaluated in Section 5.2.1. In this experiment more matches are rated as correct by using the method described above ( $\sim 41\%$ ) than by using the usual Difference-of-Gaussians approach ( $\sim 16\%$ ).

In Figure 3.5 the results of the individual steps of our interest point localization procedure are shown for the image of Figure 3.2(e). Figure 3.5(a) shows the segmentation of the hemangioma area (green border) and resulting rectangular region of interest (blue frame). For interest point localization only the green channel of this frame is used which is shown in Figure 3.5(b). Non-skin regions like the ruler in the upper left corner are ignored. Next, the Canny edge detector is applied, resulting in the edge image shown in Figure 3.5(c). Finally, interest points are detected at edge pixels having the highest gradient magnitude in a neighborhood of 3 pixels, marked as black spots in Figure 3.5(d). Thus, a total number of 1441 interest points are detected in the image.

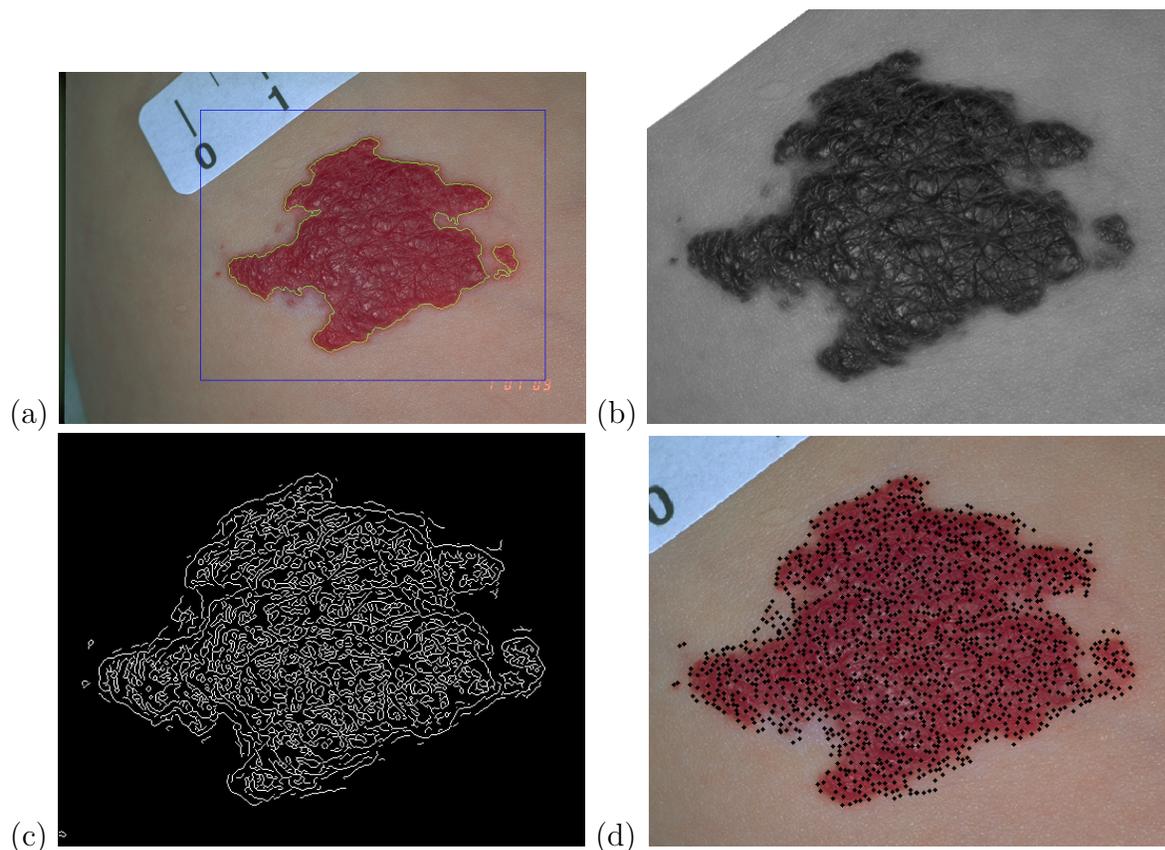


Figure 3.5: Individual steps of the interest point localization procedure.

### 3.2.2 Matching of Interest Points

Once the interest points have been localized, they have to be matched in a robust way. As in the standard SIFT method, our interest points are described by accumulating the orientations in a region around the interest point location, as illustrated in Figure 3.6. Gradient magnitudes and orientations are sampled in a  $16 \times 16$  array around the interest point location and weighted by a gaussian window, indicated by the dashed circle. The scale of all interest points is constantly set to the empirically determined value of 3.2 since no scale information is given by the edge-based interest point localization (by choosing this value a sample has a size of  $2.4 \times 2.4$  pixels). The content of each  $4 \times 4$  subregion is accumulated in an orientation histogram with 8 bins. For example, the gradients of the red square subregion produce the orientation histogram on the right side of Figure 3.6. By using a  $4 \times 4$  descriptor array of  $4 \times 4$  subregions with 8 orientation bins in each, an interest point is finally described by a 128-element feature vector.

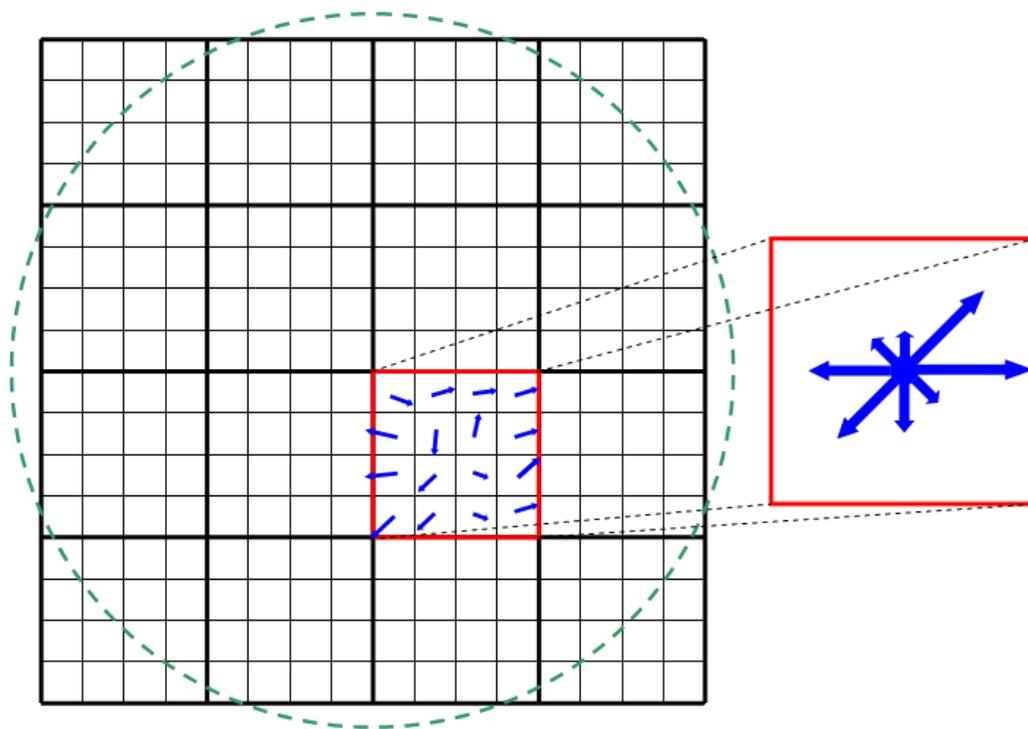


Figure 3.6: Illustration of the SIFT method for interest point description.

The task of the matching step is to find for every interest point in the sensed image the corresponding interest point in the second image. A simple way to determine the correspondences is to match interest points whose descriptor vectors have the lowest Euclidean distance to each other. However, many interest points of the sensed image might not have any corresponding interest point in the reference image and vice versa.

As described by Lowe in [25], a convenient measure for the quality of a match is not only the distance to the nearest neighbor in the reference image but also the distance to the second nearest neighbor. A “good” match for an interest point in the sensed image shows a low distance to its first nearest neighbor and a comparatively high distance to its second nearest neighbor in the reference image.

With our method the interest points are matched in a similar way. Although RANSAC is capable of handling a large portion of incorrect matches, a preselection of the matches with highest confidence can improve the stability. Therefore, matches are determined by means of lowest Euclidean distance of interest point descriptors, sorted in terms of the distance between the nearest and the second nearest neighbor, and finally only the “best”  $n$  matches are accepted. The number  $n$  of accepted matches should thereby be chosen dependent on the size of the hemangioma in the sensed image: in an image pair containing a large hemangioma more matches could be reliably detected than in an image pair containing a small hemangioma. Therefore, the value  $n$  is determined by  $n = 2\sqrt{m_{si}}$ , rounded to the nearest integer value, where  $m_{si}$  is the number of interest points detected in the sensed image. This function has been decided on empirical tests achieving the best performance on the given data. It gives enough matches for small hemangiomas (small  $m_{si}$ ) but avoids too many unreliable matches for large hemangiomas (large  $m_{si}$ ) as well. In the provided images  $m_{si}$  can vary from about 50 up to about 1500, corresponding to an  $n$ -value of 14 to 77. The function is shown in Figure 3.7 for an interest point number range of 1 to 2000.

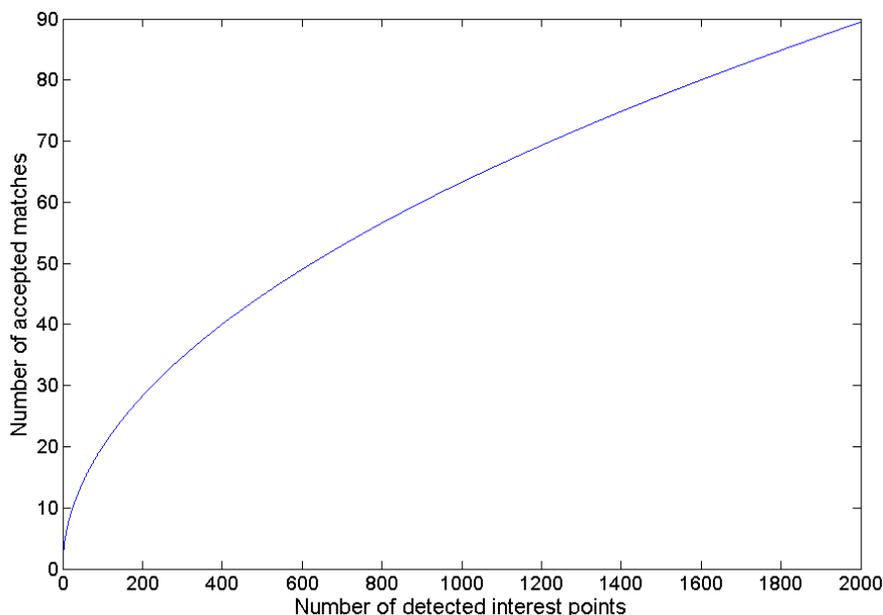


Figure 3.7: Function for determining the number  $n$  of accepted matches.

Figure 3.8 shows exemplarily the matching result between the image from Figure 3.5 and the corresponding image from the following examination. In this case 76 matches are initially accepted. However, it can be seen that still many of these matches are incorrect.

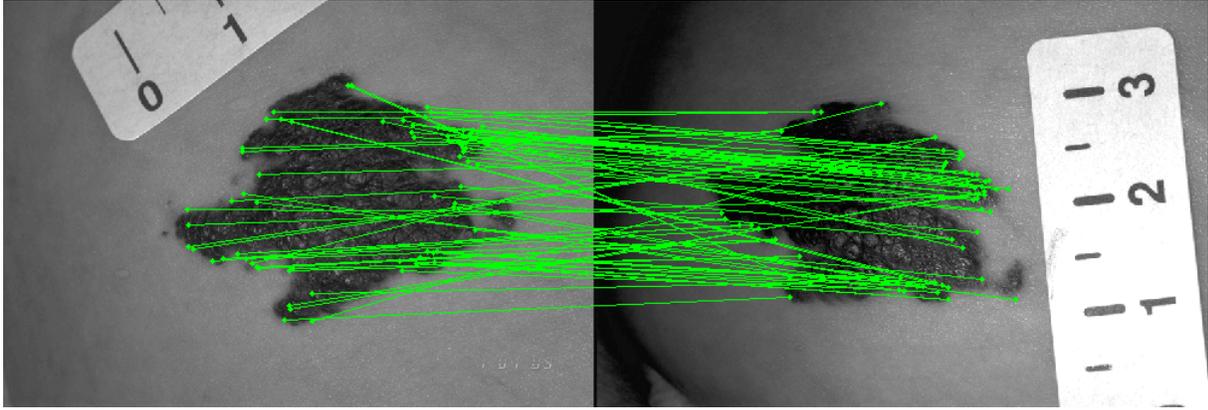


Figure 3.8: Determined corresponding interest points between two consecutive images.

The automatic exclusion of the false matches is described in the next section.

### 3.2.3 Transformation Estimation Based on Interest Point Matches

The final step in the image registration procedure is the computation of the homography that maps the sensed image onto the reference image. In general, with a homography the transformation of all points lying on a plane in the scene shown in the sensed image to the same points in the reference image can be described. This means that the hemangiomas are assumed to be planar, which is not correct in all cases. However, it is a reasonable simplification since (1) not enough correctly matched interest points are provided by the matching process and (2) matchings are not equally distributed on the hemangioma region, making non-rigid registration functions inapplicable.

This section describes the method for the robust estimation of the homography between the hemangiomas contained in the sensed and reference image. The general mathematical method for homography estimation from a set of point matches is explained below. Next, the RANSAC scheme for robust outlier detection is explained, followed by a detailed description of its application on our images.

#### Planar Homography Estimation

A homography describes the transformation between corresponding points on a plane, seen from two different perspective views, and is represented by a non-singular homogeneous  $3 \times 3$  matrix. If  $\mathbf{p}^{si} = (x^{si}, y^{si}, 1)^T$  is a point in the sensed image  $si$  and  $\mathbf{p}^{ri} = (x^{ri}, y^{ri}, 1)^T$  the corresponding point in the reference image  $ri$  lying on the same plane, their relationship is described by a homography  $H$  with

$$\mathbf{p}^{si} = H\mathbf{p}^{ri} \text{ where } H = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \quad (3.6)$$

and

$$\mathbf{p}^{ri} = H^T \mathbf{p}^{si} \quad (3.7)$$

A homography has eight degrees of freedom. Thus, it can be exactly defined by a set of four corresponding points with at most two of them being collinear. For a set of more than four matches, the homography can be estimated by the normalized Direct Linear Transform (DLT) which minimizes the algebraic error [20].

Given  $n$  matches of points  $\{\mathbf{p}_i^{si} \leftrightarrow \mathbf{p}_i^{ri}\}$  between the sensed image  $si$  and the reference image  $ri$ , a homography  $H$  is estimated by the normalized DLT in the following way:

**1. Normalization of  $\mathbf{p}_i^{si}$ :** Compute a similarity transformation  $T$  transforming the points  $\mathbf{p}_i^{si} = (x_i^{si}, y_i^{si}, 1)^T$  to a new set of points  $\tilde{\mathbf{p}}_i^{si}$  such that the centroid of the points  $\tilde{\mathbf{p}}_i^{si}$  is at  $(0, 0, 1)^T$  and their average distance is  $\sqrt{2}$ . The similarity transform  $T$  has the form

$$T = \begin{pmatrix} \sqrt{2}/\hat{d} & 0 & -\sqrt{2}\hat{x}/\hat{d} \\ 0 & \sqrt{2}/\hat{d} & -\sqrt{2}\hat{y}/\hat{d} \\ 0 & 0 & 1 \end{pmatrix} \quad (3.8)$$

with

$$\hat{x} = \frac{1}{n} \sum_{i=1}^n x_i^{si}, \quad \hat{y} = \frac{1}{n} \sum_{i=1}^n y_i^{si}, \quad \hat{d} = \frac{1}{n} \sum_{i=1}^n \sqrt{(x_i^{si} - \hat{x})^2 + (y_i^{si} - \hat{y})^2} \quad (3.9)$$

**2. Normalization of  $\mathbf{p}_i^{ri}$ :** Likewise, compute a similarity transformation  $T'$  transforming  $\mathbf{p}_i^{ri}$  to  $\tilde{\mathbf{p}}_i^{ri}$ .

**3. Direct Linear Transformation:** For each match  $\{\tilde{\mathbf{p}}_i^{si} \leftrightarrow \tilde{\mathbf{p}}_i^{ri}\}$  the homography  $\tilde{H}$  is given by  $\tilde{\mathbf{p}}_i^{ri} = \tilde{H} \tilde{\mathbf{p}}_i^{si}$ . This can be written in the form

$$A_i \mathbf{h} = \begin{pmatrix} \mathbf{0}^T & -\tilde{\mathbf{p}}_i^{siT} & y_i^{ri} \tilde{\mathbf{p}}_i^{siT} \\ -\tilde{\mathbf{p}}_i^{siT} & \mathbf{0}^T & -\tilde{x}_i^{ri} \tilde{\mathbf{p}}_i^{siT} \\ -\tilde{y}_i^{ri} \tilde{\mathbf{p}}_i^{siT} & \tilde{x}_i^{ri} \tilde{\mathbf{p}}_i^{siT} & \mathbf{0}^T \end{pmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = \mathbf{0} \quad (3.10)$$

$A_i$  is a  $3 \times 9$  matrix and  $\mathbf{h}$  is a 9-element vector made up of the entries of the matrix  $\tilde{H}$ ,

$$\mathbf{h} = \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix}, \quad \tilde{H} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \quad (3.11)$$

with  $h_i$  the  $i$ -th element of  $\mathbf{h}$ . Since only two rows of  $A_i$  are linearly independent, each match gives two equations in the entries of  $\tilde{H}$ . We therefore use only the first two rows

of  $A_i$  and assemble the  $n$   $2 \times 9$  matrices into a single  $(2n) \times 9$  matrix  $A$ . The solution  $\mathbf{h}$  is the unit singular vector corresponding to the smallest singular value of  $A$  obtained by a Singular Value Decomposition.

**4. Denormalization:** Denormalize  $\tilde{H}$  with  $H = T'^{-1}\tilde{H}T$ .

To transform an image point  $\mathbf{p}_i^{s_i}$  of the sensed image it has to be represented by homogeneous coordinates  $\mathbf{p}_i^{s_i} = (x_i^{s_i}, y_i^{s_i}, 1)^T$ . The transformation is achieved by  $\mathbf{p}_i^{r_i} = H\mathbf{p}_i^{s_i}$  and division of the resulting point  $\mathbf{p}_i^{r_i}$  by its homogeneous component.

### Random Sample Consensus

RANSAC (RANdom SAMple Consensus) is a robust model fitting method which is able to cope with a large portion of incorrect data samples [11]. Its underlying assumption is that in a set of observations the number of correct samples (the *inliers*) being consistent with the correct model outnumbers the number of incorrect samples (the *outliers*) being consistent with an incorrect model because the outliers are randomly distributed.

The RANSAC scheme works as follows: assume that we have  $M$  samples from an observation and  $N$  samples are needed to determine the parameters of the model. Then the following steps are repeated until the performance goal is met or the maximum number of iterations is reached:

1. Randomly choose a subset of  $N$  samples out of the set of  $M$  samples and estimate the parameters of the model.
2. Determine the inliers of this model: number of samples fitting the model within user-defined tolerance.
3. Choose this model if it shows the maximum number of inliers found so far.

Optionally, in the end the model's parameters can be re-estimated with all inliers.

### Homography Estimation between Hemangioma Images

As mentioned above, the matching of interest points also produces false matches that have to be detected and discarded. Since every match is equally considered for homography estimation by the DLT algorithm, it is not robust against these so called outliers, i.e. even one outlier can disturb the registration result to a high degree. Therefore, a necessary requirement of the final homography estimation method is a robust detection of inliers and outliers in the present interest point matches.

For that purpose the RANSAC scheme is applied to our matched interest points. In our case the number  $M$  of total samples is the number of initially detected matches (see Section 3.2.2). Our model is a homography and the parameters to be estimated are the 9 elements of the  $3 \times 3$  homography matrix. For homography estimation 4 matches are needed, i.e.  $N = 4$ . An inlier in our case is defined as a match where the interest point in the sensed image is located within a given distance to the matching interest point in the reference image after a transformation with the homography.

On our images best results were achieved by iterating 2000 times and allowing a maximum distance of 5 pixels for the inlier decision (in a typical image 5 pixels correspond to  $\sim 0.375$  mm). To increase robustness only homographies with an absolute value of the determinant in the range of 0.1 to 10 are allowed. If the determinant of a homography or its inverse is close to zero, it corresponds to a degenerate case. The full algorithm is summarized in Algorithm 1.

Once the final homography is computed, the last step in the registration process is to transform the sensed image onto the reference image. In practice, this is done by computing the inverse of the homography, transforming each image point of the new image onto the sensed image and computing its image value by bilinear interpolation. In Figure 3.9 the final results for the images of Figure 3.2(e) and Figure 3.2(f) are shown. Figure 3.9(b) shows the remaining inliers determined by RANSAC of the initial matches depicted in Figure 3.9(a). In this example 31 of the 76 initial matches are classified as inliers. The final transformed sensed image can be seen in Figure 3.9(c). The difference between the transformed image and the reference image is shown in Figure 3.9(d).

```

for  $i=1$  to 2000 do
  ▷ select a random sample of 4 matches out of all matches and compute the
  homography  $H$ 
  if (absolute value of the determinant of  $H$  between 0.1 and 10) then
    ▷ calculate the distance  $d$  to all other matches
    ▷ compute the number of inliers as the number of matches for which  $d < 5$ 
    if ( $H$  has number of inliers larger than the present best homography) then
      ▷ mark  $H$  as best homography
    else if (number of inliers is equal) then
      ▷ choose the homography with the lower standard deviation of inliers
    end
  end
end
end
  ▷ re-estimate the best homography from all matches classified as inliers using the
  normalized DLT

```

**Algorithm 1:** The RANSAC-algorithm for estimating a homography from putative matches between two hemangioma images.

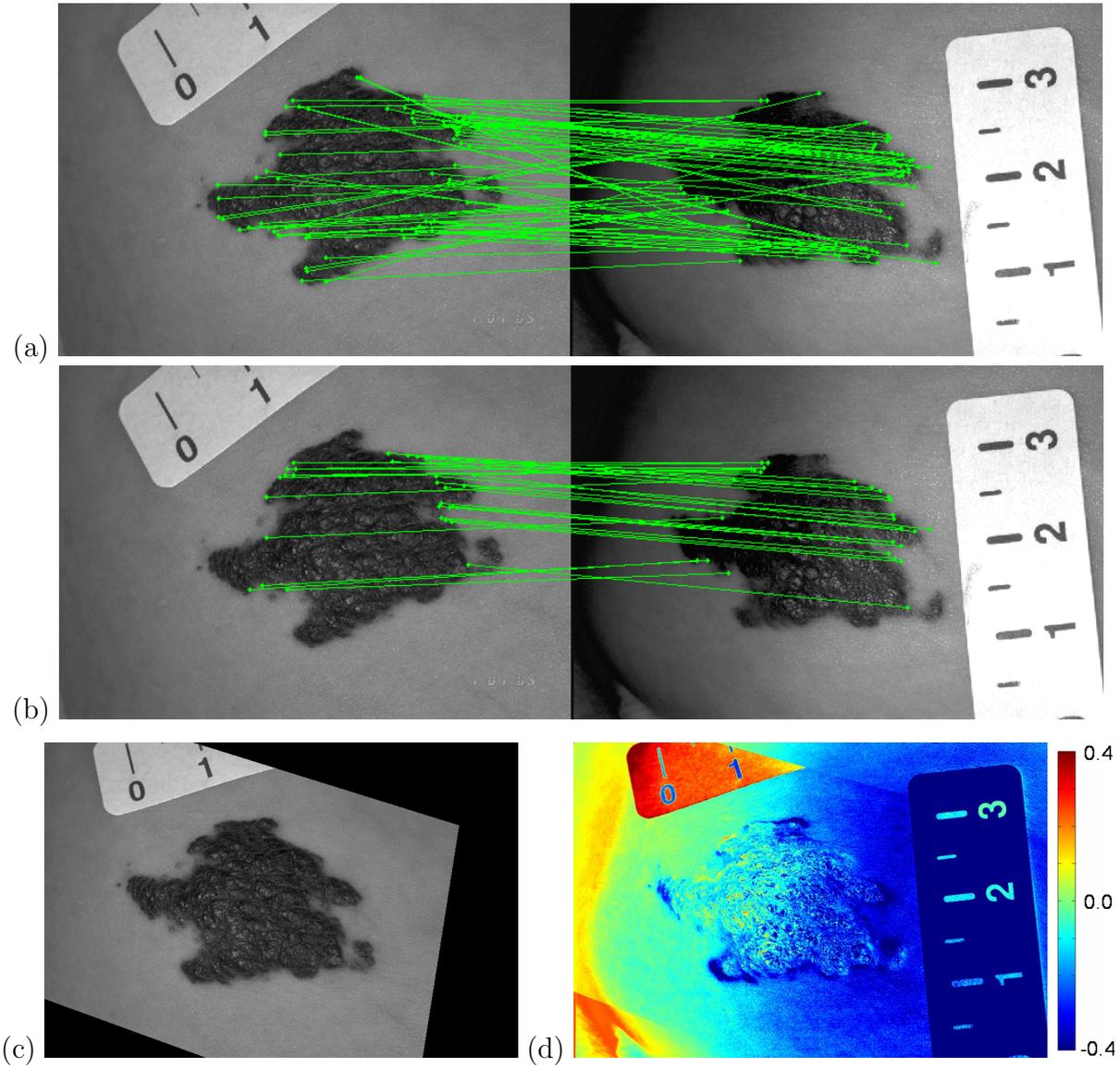


Figure 3.9: (a) Initially detected matches between sensed image (left) and reference image (right), (b) detected inliers, (c) transformed sensed image and (d) image displaying the difference between the transformed sensed image and the reference image.

# Chapter 4

## Detection of Regressing Hemangioma Regions

This chapter describes the process for the detection of regressing hemangioma parts in follow-up images, using the hemangioma segmentation as region of interest and color differences to the first examination for a pixel-wise classification of the hemangioma region. In the beginning hemangiomas typically exhibit a saturated red color, whereas during the course of therapy regressions appear as pale gray regions, a process often referred to as *graying* [6]. Usually, these regressions do not occur uniformly on the whole hemangioma, but start at certain regions and expand during time [6]. According to dermatologists, the detection and sizing of such regions is an important parameter to assess the healing process.

As an example, see the two follow-up images in Figure 4.1(a) and Figure 4.1(b). The image of Figure 4.1(a) was taken at the first examination and the whole hemangioma is reddish saturated, expect for a small inner region. Image Figure 4.1(b) was taken 6 weeks later and yet the hemangioma shows regressions in its inner regions. These regressions are marked with a white border in Figure 4.1(c).



Figure 4.1: (a),(b) Consecutive images of a hemangioma and (c) regressions marked with white border.

Since the detection of regressing regions relies on their color, a pixel-wise classification scheme similar to the classification step in hemangioma segmentation (Section 2.2.3) is applied to a single image. In addition to the color features, the first examination is taken into account to achieve a more specific determination of the change due to the regressing hemangioma for each pixel. Motivated by the hemangioma's typical healing process de-

scribed above, it is assumed that the hemangioma has a red saturated color at the first examination and that therefore the local color change provides a strong feature for the classifier. For this purpose the registration introduced in Chapter 3 provides the necessary correspondence between the current image and the first reference image.

The whole procedure of the proposed method for the detection of regressing hemangioma regions is depicted in Figure 4.2: consider a series of four follow-up images (Image 1 to Image 4). Between each consecutive image a homography was computed by the registration method described in Section 3.2 ( $H_{12}$ ,  $H_{23}$  and  $H_{34}$ ). In the first step the region encompassing the hemangioma is determined by the segmentation method of Section 2.2 (white area). Next, the mask defining the hemangioma region is transformed with the composite homography  $H = H_{34}H_{23}H_{12}$  to define the region of interest for the classification process. In the last step, in this region of interest all pixels are classified in *regressing* (gray area in Figure 4.2) and *not regressing*, using also the differences in color values between Image 1 and Image 4 computed via the composite homography  $H$  (this is indicated by the double-headed arrow with the minus sign in Figure 4.2).

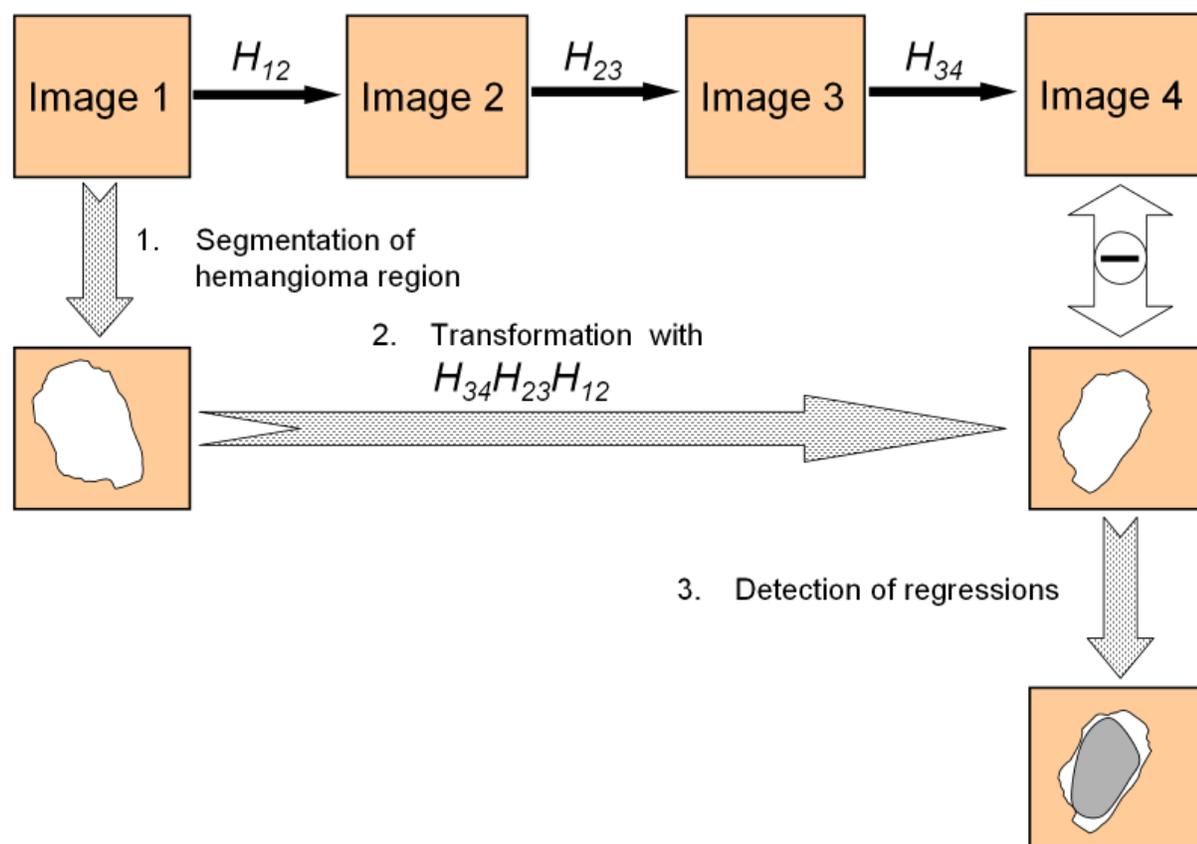


Figure 4.2: Illustration of the regression detection procedure.

The pixel-wise classification is addressed in Section 4.1, including a description of the classifier used and the selection of appropriate features. A detailed step-by-step explanation

of the whole methodology for regression detection is given in Section 4.2.

## 4.1 Pixel-wise Classification

Similar to the segmentation of saturated hemangioma regions (Section 2.2), a pixel-wise classification scheme based on color features is used for the regression detection task. Thereby, the hemangioma region is classified into parts showing a regression and parts showing no or minor changes compared to the first examination. By using the registration method presented in Section 3.2, classification does not have to rely solely on the current color features, but can also use the color information from the image of the first examination. The change of colors between the sessions is an important and discriminative feature, which will be shown in Section 5.3.3 later on.

The images in Figure 4.1 point out that the classification of a hemangioma into regressing and not regressing regions is less distinct than for the hemangioma segmentation method presented in Chapter 2, since color differences are considerably lower. On this account, a more sophisticated classifier, the Neural Network [2], is applied to the data. This type of classifier is introduced in Section 4.1.1. The selection of an appropriate set of features for regression detection is described in Section 4.1.2.

### 4.1.1 Neural Networks

Neural Networks [2] are extensions to the simple perceptron introduced in Section 2.2.3. Whereas a perceptron consists only of one input layer and output layer, one or more hidden layers of units between the input and output layer are added to a neural network structure. Thus, in literature neural networks are often also referred to as *multi-layer perceptrons* [2]. Figure 4.3 shows a simple example for the structure of a neural network. The network consists of 4 input units with input values  $x_i$ , 1 hidden layer with 3 hidden units having output values  $v_j$  and 2 output units with output values  $o_k$ .

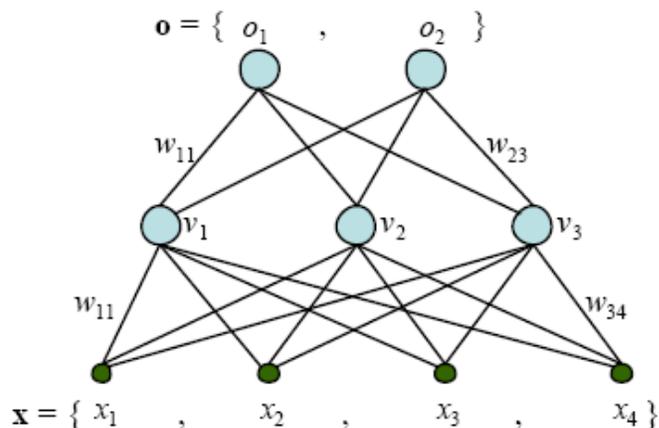


Figure 4.3: Example for a neural network structure.

In the neural network each unit of a layer is connected to each unit of the next layer and each connection possesses a weight  $\omega_{ij}$ . Additionally, for each unit a bias  $\theta$  and a differentiable activation function (e.g. the *sigmoid* or *tanh* function [2]) is defined. The output of a given network is computed as follows: The weights are multiplied with the corresponding input values and summed up. To this sum the bias of the unit is added and the actual output value  $\nu_j$  is determined by the output of the activation function. These output values serve as input values for the next units/layer and so on. Thus, the state of a certain network is defined by its weights and biases. This state has to be trained by means of a training set with given output values.

## Neural Network Training

For neural network training the *backpropagation* algorithm [2] is applied. With backpropagation, the following steps are conducted until the network error is small enough or the maximum number of epochs is reached:

1. **Forward-Pass:** Computation of the network output  $\mathbf{o}$  by a certain input vector  $\mathbf{x}$ .
2. **Determination of network error:** The network output  $\mathbf{o}$  is compared to the correct output  $\mathbf{t}$  and the error is computed by means of a certain error metric. For instance, the output can be assessed by the *mean squared error*:  $mse = \frac{1}{N_k} \sum_{k=1}^{N_k} (o_k - t_k)^2$ , where  $N_k$  is the number of output units.
3. **Backward-Pass:** The weights are successively adapted on the basis of the network error, starting from the output layer in direction to the input layer. The weights and the error metric chosen define an error surface that has to be minimized by adjustment of the weights. This can be accomplished by miscellaneous optimization methods, e.g. gradient descent or quasi-newton. A detailed discussion of the various methods is out of the scope of this thesis, for an overview see [1].

## Regularization by Weight Decay

The task of a neural network is not the exact representation of the training data but rather the building of a statistical model that describes the data as well as possible. In other words, the network should exhibit a good *generalization*, that is, to make good predictions for new inputs. One way to avoid so-called *over-fitting* of the network is *regularization* which adds a penalty  $\Omega$  to the error function. The error of a network is then given by

$$error = \gamma \cdot mse + (1 - \gamma) \cdot \Omega \quad (4.1)$$

where  $\gamma$  is called the *performance ratio* defining the extent to which the penalty  $\Omega$  influences the error. A simple form of the penalty term  $\Omega$  is called *weight decay*, given by

$$\Omega = \frac{1}{N} \sum_{i=1}^N \omega_i^2 \quad (4.2)$$

where  $N$  is the number of weights in the network. This type of penalty prevents the network from producing extreme weights overfitting the training data.

### 4.1.2 Feature Selection

As mentioned above, color features are used for classification. For this reason, the individual components of the three color spaces RGB, HSV and L\*a\*b\* along with the feature *abdist* were tested for their suitability (for a description of these features see Section 2.2.3). The change of colors is included in the classification procedure by computing the difference of the color features between the transformed sensed image and the reference image, termed *difference features* in the following. Thus, along with difference features in total 20 features for classification are available. To deal with varying illumination conditions during image acquisition all features are skin-normalized by the method presented in Section 2.2.2. Texture features (local entropy, local range and standard deviation [15]) were also tested but abandoned, because they did not provide relevant information for the classifier during initial experiments.

To enhance the classifier's performance and classification time, irrelevant and redundant features have to be removed from the total set of features. A first subset of features was achieved by applying the feature selection method proposed by Krizek et al. [24]. Their algorithm removes irrelevant and redundant features by a weight modification of training samples similar to the AdaBoost algorithm [37]. Features are iteratively selected on the basis of their discriminative power measured by a simple error function and at every iteration step weights of training samples are modified so that the currently selected feature appears irrelevant w.r.t. new weights, thus allowing for a more proper determination of the discriminative power of remaining features.

Using this method on our data, the following features were selected out of the total 20 features:

- Color features
  - $G$  and  $B$  from RGB
  - $S$  from HSV
  - $a^*$  from L\*a\*b\*
  - *abdist*
- Difference feature
  - $G$  and  $B$  from RGB
  - $L^*$  and  $a^*$  from L\*a\*b\*

However, further experiments have shown that there is still redundancy in this feature subset and that the performance of the neural network is not altered by a removal of certain features. Thus, in the end the following four features were selected:

- Color features
  - $G$  from RGB
  - $a^*$  from L\*a\*b\*
  - *abdist*

- Difference feature
  - $G$  from RGB

This choice is plausible, since these color features are also used for the hemangioma segmentation (see Section 2.2.3). In Figure 4.4 the values of these chosen features are exemplarily shown for the image of Figure 4.1(c), normalized to the range 0 to 1 for presentation (black represents 0, white represents 1).

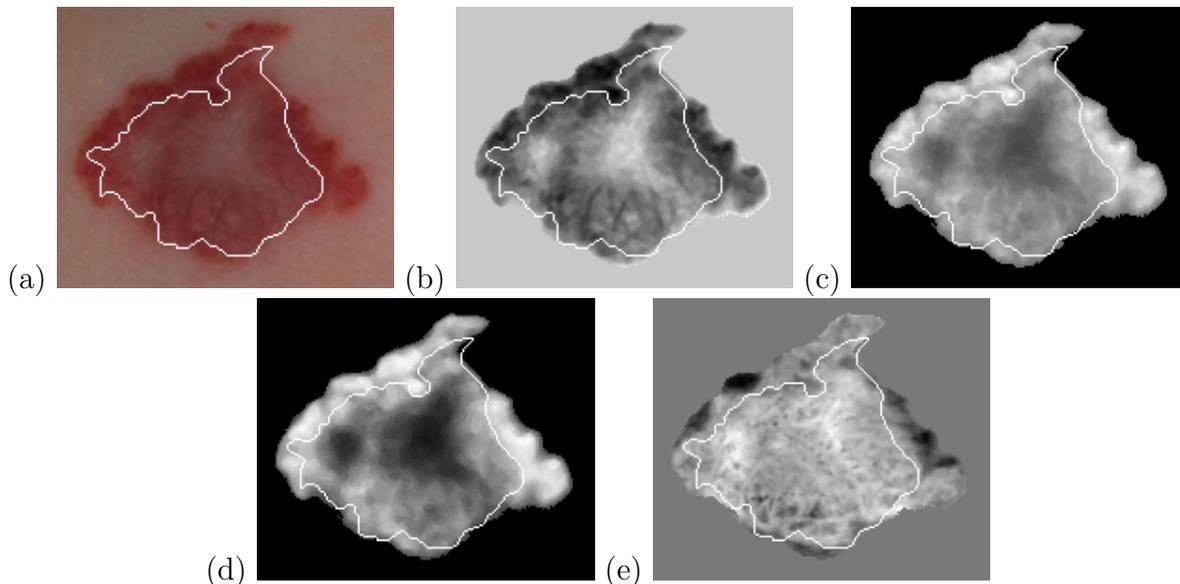


Figure 4.4: (a) Original hemangioma image with regressing region marked by white border and values of features: (b)  $G$ , (c)  $a^*$ , (d)  $abdist$ , (e) difference of  $G$  between first follow-up and current image. Values are normalized to the range 0 to 1, thus 0 is represented by black and 1 is represented by white.

Apparently, differences between regressing and not regressing regions are less distinctive than for the segmentation of the hemangioma area (see Figure 2.6 for a comparison). Nevertheless, the experiments in Section 5.3.2 show that the method achieves a similar accuracy.

## 4.2 Regression Detection in Registered Hemangioma Follow-up Images

In this section the whole process of regression detection in follow-up series is described in detail. Consistent with the illustration shown in Figure 4.2, the following steps were performed for regression detection in the  $i$ th image of a hemangioma follow-up series:

1. **Segmentation of hemangioma region in the first image:** The hemangioma region in the first image is determined by the segmentation algorithm presented in

Chapter 2. This is needed to define a region of interest for the pixel-wise classification. Although the pixel-wise detection of regressions could be performed on the whole image as well, classification performance is increased since only relevant data has to be classified into *regressing* and *not regressing*, i.e. no healthy skin is involved.

2. **Registration of first image with  $i$ th image:** The first image is registered with  $i$ th image, i.e. the homography  $H$  between these two images is computed. For that purpose the registration algorithm presented in Chapter 3 is used. For robustness,  $H$  is not determined directly by a registration of the first image with the  $i$ th image, but by a concatenated registration of the intermediate images: at first the homography  $H_{1,2}$  between the first and the second image is determined, next the homography  $H_{2,3}$  between the second and the third image, and so on until the homography  $H_{i-1,i}$ . The composite homography is then simply computed by matrix multiplication:  $H = H_{i-1,i} \dots H_{2,3} H_{1,2}$ . Since the changes in hemangioma appearance between the first and the  $i$ th image increase more and more during follow-up, this scheme of concatenation is more robust than a direct registration of the first and  $i$ th image.

The need of the homography  $H$  is twofold: Firstly, it is used to transform the segmentation of the first image to define the region of interest in the  $i$ th image. Naturally, this could also be accomplished by hemangioma segmentation in the  $i$ th image, but the segmentation of the first image is more robust and accurate, since at the first examination the hemangioma shows less regressions and therefore a comparatively high contrast between skin and hemangioma. And secondly, it is needed to build difference features for the pixel-wise classification.

3. **Feature extraction:** After the homography between the first and  $i$ th image is determined, the color features for classification are extracted from the images (see Section 4.1.2). Due to various illumination conditions during image acquisition, both the first and the  $i$ th image are normalized w.r.t. their skin color by the same method used for hemangioma segmentation (see Section 2.2.2). The features  $G$ ,  $a^*$  and  $abdist$  are then directly extracted from the  $i$ th image. The color difference of the green band in RGB space ( $G$ ) is computed by image subtraction of the first image, transformed by  $H$ , with the  $i$ th image.
4. **Pixel-wise classification:** The extracted features are used to classify all pixels in the region of interest in *regressing* and *not regressing*, using the neural network classifier described in Section 4.1.1.

# Chapter 5

## Experiments

This chapter deals with the evaluation of the methods presented in the thesis. In Section 5.1 the segmentation algorithm described in Chapter 2 is evaluated. Experiments regarding the registration method proposed in Chapter 3 are reported in Section 5.2. The regression detection described in Chapter 4 is evaluated in Section 5.3. Detailed information about the experiments conducted is given in the respective sections.

Each description is divided into *Setup*, where the conducted experiment is described in detail, and *Results*, where the corresponding results are briefly presented. A detailed overall discussion about the obtained results and the implications to practice are finally given in Section 5.4.

### 5.1 Measurement of Hemangioma Size

The evaluation of the proposed method for hemangioma surveying (Chapter 2) is done in various tests: first of all the precision of the scale computing algorithm is measured in Section 5.1.1. Second, in Section 5.1.2 the area measurement is conducted on a set of synthetic images. Third, the segmentation of hemangioma is evaluated by tests on 120 clinical images manually annotated under the supervision of a dermatologist (Section 5.1.3), and fourth, the accuracy of the whole surveying process is estimated by a comparison of 20 image pairs, each showing the same hemangioma at the same time of acquisition, i.e. the areas should be equal if perfect measurement would have been achieved (Section 5.1.4).

#### 5.1.1 Accuracy of the Scale Computing Algorithm

**Setup:** For this experiment we computed the Euclidean distance between two marks of the ruler of 20 images with the algorithm proposed in Section 2.1 and compared it with manually determined values (Table 5.1). These values are achieved by manually choosing the longest Euclidean distance between two marks of the ruler visible in the image. Nevertheless, these values do not necessarily represent the real spatial resolution of the image since we do not take into account the curvature and position of the ruler just as in our algorithm. In addition the manually determined Euclidean distance is also not exact, a failure of 1-2 pixels is possible. However, the experiments reveal the ability of the

algorithm to consistently determine the scale w.r.t. the standard of reference annotation on the images. The effect of image acquisition on the measurement is captured by the experiments in Section 5.1.4.

Image Pair	Computed Euclidean Distance (pixels)	Manually Determined Euclidean Distance (pixels)	Error (%)
1	130	129	0.78
2	127	124	2.42
3	132	129	2.33
4	136	134	1.49
5	140	138	1.45
6	135	133	1.50
7	139	135	2.96
8	139	139	0.00
9	130	130	0.00
10	133	131	1.53
11	134	141	4.96
12	130	129	0.78
13	145	145	0.00
14	141	151	6.62
15	98	100	2.00
16	140	141	0.71
17	146	146	0.00
18	130	132	1.52
19	141	141	0.00
20	146	146	0.00
Average			<b>1.55</b>

Table 5.1: Comparison between the computed and the manually determined Euclidean distance between two marks of the ruler in 20 images.

**Results:** As can be seen in Table 5.1 the average error made by the method is low (average error rate 1.55%). Larger errors of more than 3 % arise only on image 11 and 14. These are shown in Figure 5.1. In image 11 (Figure 5.1(a)) there can be seen a white cloth that is segmented together with the ruler. Therefore, in the computed rotated ruler mask the three scanlines (red) are too far below and erroneously the distance between the numbers instead of the marks is taken for scale computing (Figure 5.1(b)). The same problem occurs in image 14 (Figure 5.1(c) and Figure 5.1(d)) but this time caused by the fact that only about the half of the ruler is visible in the image.

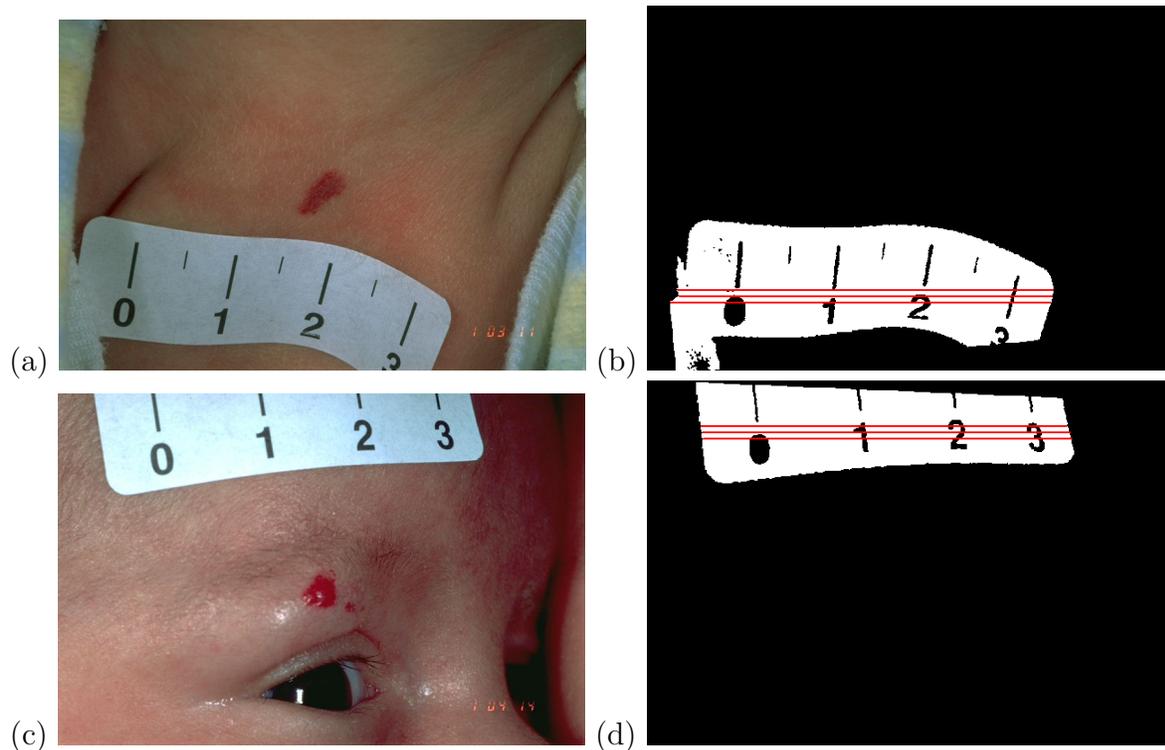


Figure 5.1: Image 11 and 14 of Table 5.1 with corresponding computed ruler masks. The three scanlines are marked red.

### 5.1.2 Accuracy of Hemangioma Area Measurement on Synthetic Images

**Setup:** The measurement of the hemangioma area was tested on a set of 10 synthetically generated images. In such images the error in area measurement is not influenced by geometric distortions and illumination inconsistencies, thus giving us an estimate of the algorithm's performance under optimal conditions. In total, 10 images were generated in the following way: an arbitrary shape representing the hemangioma was painted with the typical hemangioma color (red portion = 100, green portion = 0, blue portion = 15 of 255) and placed in a skin image along with a randomly rotated ruler. Additionally, gaussian noise with mean  $\mu = 0$  and variance  $\sigma^2 = 0.5$  was added and the images were smoothed with a  $7 \times 7$  gaussian filter with standard deviation  $\sigma = 1.0$  to simulate the smooth transition between skin and hemangioma along the border. In Figure 5.2 three such synthetic images with detected segmentations marked by a green border are shown. For the experiments the area of the hemangioma measured with our algorithm is compared to the real known area and the relative errors were recorded.

**Results:** The results can be seen in Table 5.2. The average relative error in area measurement is 0.62 %. The scale of the images is correctly measured in all 10 images, thus the error is only caused by the uncertainty of correct hemangioma border due to the smoothing with the gaussian filter. Generally, relative errors are higher on images with smaller

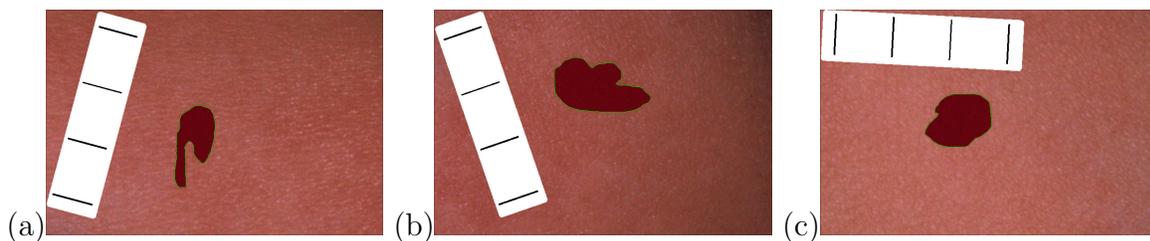


Figure 5.2: Three synthetic hemangioma images with the automatic segmentation marked by a green border.

Image	Real Area ( $cm^2$ )	Measured Area ( $cm^2$ )	Error (%)
1	0.9372	0.9404	0.35
2	0.6013	0.6037	0.40
3	0.4107	0.4136	0.69
4	0.6140	0.6168	0.46
5	1.1407	1.1454	0.41
6	0.7782	0.7814	0.42
7	0.0404	0.0392	2.83
8	0.5108	0.5083	0.48
9	1.3217	1.3234	0.13
10	0.8449	0.8456	0.09
Average			<b>0.62</b>

Table 5.2: Results of area measurement on 10 synthetic hemangioma images.

hemangiomas than on images with larger hemangiomas, since smaller hemangiomas show a higher portion of border pixels.

### 5.1.3 Accuracy of the Hemangioma Segmentation Algorithm

**Setup:** Experiments were performed on a set of 120 images gathered during clinical examinations. 30 images were used for training and the remaining 90 images served as the test set. For all images the relative ground truth was obtained by a manual segmentation in a region of interest under the supervision of a medical expert. To evaluate the accuracy of the segmentation for every image the following error metrics were applied:

- error rate =  $\frac{\text{number of misclassified pixels}}{\text{total number of pixels}}$
- false positive rate =  $\frac{\text{nr. of false positive pixels}}{\text{nr. of negative pixels}}$
- false negative rate =  $\frac{\text{nr. of false negative pixels}}{\text{nr. of positive pixels}}$
- absolute area difference =  $|Area(A) - Area(M)|$
- border error =  $\frac{Area(A \cup M) - Area(A \cap M)}{Area(M)}$

where  $A$  and  $M$  are the regions obtained by the automatic segmentation and the manual segmentation, respectively. The proposed formula for the border error (adopted from [19]) is the most significant error metric because it is independent of both the size of the hemangioma and the size of the region of interest.

**Results:** The average border error on the 90 test images is 32.1 %, as can be seen in Table 5.3 where the average errors of all 90 test images are shown. The average error made in the surveying of the hemangiomas lies at  $0.0965 \text{ cm}^2$ , where the average hemangioma size is  $0.6132 \text{ cm}^2$ . Generally, the obtained average border error is hardly influenced by few particular outliers with border errors of more than 100 %. Table 5.4 shows the distribution of the different border errors of all 90 images. It can be seen that the majority of the images (54 of 90) could be segmented with a border error of less than 20 % and only 15 images yield an error of more than 50 %.

In Figure 5.3 some results are depicted. Figure 5.3a-c belongs to the best segmentation results with border errors of 3.6 %, 5.7 % and 6.8 %, respectively. Figure 5.3d-f belongs to the worst segmentation results with border errors of 247.7 %, 137.5 % and 141.2 %, respectively.

#### 5.1.4 Precision of Automatic Measurement of Hemangioma Size

**Setup:** To assess the precision of the entire procedure of hemangioma size measurement the hemangioma area on 20 pairs of images, depicting the same hemangioma and taken within a few minutes, were measured. In the absence of error, both images should have precisely the same computed hemangioma area. The actual error is estimated by the absolute area difference and variation coefficient of both measurements. The variation coefficient of multiple measurements is defined as the their standard deviation divided by their mean. To have an estimate of the minimal achievable error, given the differences in image acquisition reference measurements with manual segmentation and scale computing were made on the same image pairs.

**Results:** The results of both the automatic and the manual measurements of all 20 image pairs are listed in Table 5.5. The average difference of hemangioma area is  $0.0775 \text{ cm}^2$  while the average variation coefficient is 8.82 %. This error is mainly caused by variations

Error Rate	False Pos. Rate	False Neg. Rate	Absolute Area Difference	Border Error
6.8 %	5.5 %	11.6 %	$0.0965 \text{ cm}^2$	32.1 %

Table 5.3: Mean errors on 90 test images.

Border Error	0-20 %	20-50 %	> 50 %
Number of Images	54	21	15

Table 5.4: Distribution of border error on 90 test images.

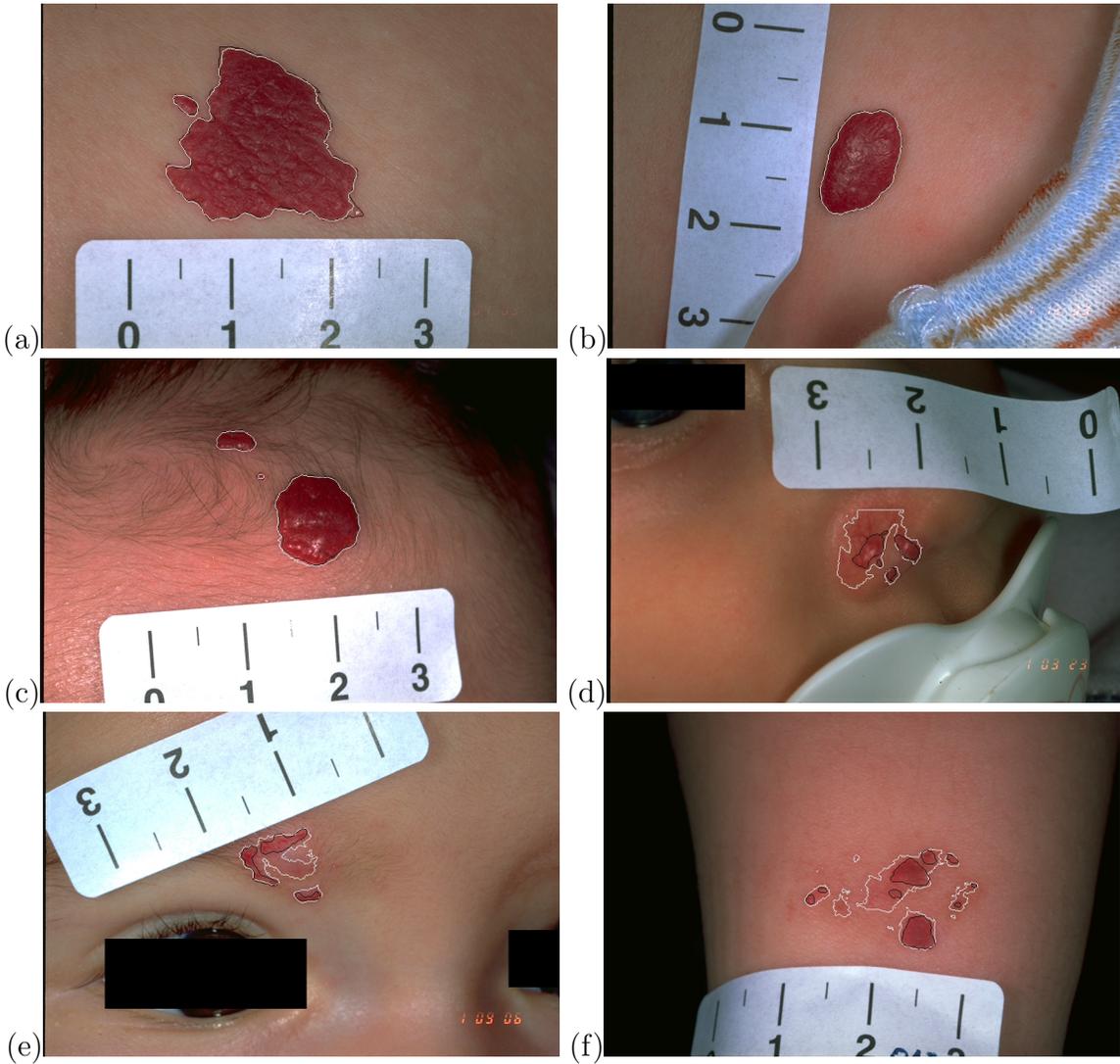


Figure 5.3: Automatic segmentation (white) and ground truth (black) of 6 images.

in the image acquisition procedure since geometric distortions caused by an improper camera viewpoint are not considered by the method. The reference measurements with manual segmentation and scale computing leads to an average difference of  $0.0394 \text{ cm}^2$  and an average variation coefficient of 4.53 %.

## 5.2 Registration of Follow-Up Hemangioma Images

This section deals with the evaluation of the method for the registration of hemangioma follow-up images presented in Chapter 3. At first, a comparative evaluation of Canny edge interest points and standard Difference-of-Gaussians interest points is given in Section 5.2.1. For an evaluation of the image registration method under “perfect” conditions with unchanged hemangioma appearances experiments were conducted on a set of 20

Image Pair	Difference Manual Measurement ( $cm^2$ )	Variation Coefficient Manual Measurement (%)	Difference Automatic Measurement ( $cm^2$ )	Variation Coefficient Automatic Measurement (%)
1	0.1727	4.58	0.2290	6.45
2	0.0245	7.81	0.0340	11.92
3	0.0277	5.79	0.0155	3.09
4	0.0381	0.64	0.1543	2.97
5	0.0002	0.04	0.0071	2.30
6	0.0129	5.56	0.0225	17.73
7	0.0064	0.70	0.0430	4.96
8	0.1027	3.31	0.2302	8.12
9	0.0027	1.54	0.0023	1.29
10	0.0736	8.38	0.0447	5.78
11	0.0066	2.41	0.0212	8.04
12	0.0139	7.65	0.0243	12.76
13	0.0015	0.03	0.4261	11.21
14	0.0681	5.89	0.1037	9.32
15	0.0109	6.40	0.0783	35.92
16	0.0089	3.56	0.0150	10.70
17	0.0243	8.27	0.0129	4.35
18	0.0647	1.98	0.0042	0.15
19	0.1183	11.93	0.0487	5.43
20	0.0087	4.11	0.0325	16.82
Average	0.0394	4.53	<b>0.0775</b>	<b>8.82</b>

Table 5.5: Differences in manually and automatically determined areas between two images of the same hemangioma.

image pairs, each pair showing the same hemangioma at the same time, i.e. only a few seconds between image acquisitions. The results are shown and discussed in Section 5.2.2. In order to assess the amount of deterioration for registrations of consecutive images (i.e. with a period of several weeks between image acquisitions), tests on four image series are reported in Section 5.2.3.

### 5.2.1 Comparison of Difference-of-Gaussians and Canny Edge Interest Points

**Setup:** As described in Section 3.2.1 interest points are detected at Canny edge points having a local maximum of gradient magnitudes. In this section the improvement by means of reliable matches using these points instead of the original Difference-Of-Gaussians interest points is demonstrated by a simple test where we have compared the number of correct matches of three image pairs of consecutive hemangioma images. For every image pair the best 40 matches having the lowest Euclidean distance are determined and the correct matches are counted for both methods by visual inspection.

**Results:** As can be seen in Table 5.6, with our localization method  $\sim 41\%$  of the matches are rated as correct whereas with the DoG localization method (using the standard parameter values defined in [25]) only  $\sim 16\%$  are rated as correct. As a conclusion, on our images Canny interest points are much more stable than Difference-of-Gaussians interest points.

Image Pair	Correct Matches DoG	Correct Matches Canny
1	7/40	16/40
2	7/40	21/40
3	5/40	12/40
Total	19/120	49/120
Percent	$\sim 16\%$	$\sim 41\%$

Table 5.6: Comparison of Difference-of-Gaussians interest points and Canny edge interest points by manually determined correct matches.

### 5.2.2 Accuracy of Registration on Hemangioma Images taken at the same Time

**Setup:** In order to increase reliability two or more images of a hemangioma were taken during an examination. 20 of such image pairs are used for testing the precision of the proposed image registration method. The registration error is measured using three different metrics:

1. **Distance Error of Inliers:** The average pixel distance of transformed inliers to the real points of the matches classified as inliers.
2. **Distance Error of 5 Test Points:** For each image pair 5 matches are manually placed and the average pixel distance achieved with the estimated homography is measured.
3. **Border Error:** Both the sensed and reference image are segmented and the border error (see Section 5.1.3) between the transformed sensed segmentation and the reference segmentation is measured.

**Results:** The results of the test are listed in Table 5.7. Additionally to the proposed error metrics for each image pair the number of detected interest points in the sensed and reference image, the resultant number of initial matches and the fraction of these matches classified as inliers are reported. It can be seen that the average distance error of the 5 test points lies at 2.31 pixels ( $\sim 0.175$  mm) which is not a large increase compared to the distance error of the inliers (1.61 pixels or  $\sim 0.125$  mm). Another indication of accuracy is the fraction of the initial matches classified as inliers. On average 35.3 of the 36.7 initial

Image Pair	Interest points sensed image	Interest points reference image	Initial Matches	Inliers	Average Distance Error of Inliers (pixels)	Average Distance Error of 5 Test Points (pixels)	Border Error (%)
1	590	659	49	48	1.69	3.28	10.78
2	391	489	40	40	0.55	1.51	7.85
3	1441	1343	76	76	0.73	1.63	1.76
4	142	84	24	20	2.98	2.74	8.93
5	225	294	30	30	0.70	2.41	3.30
6	393	360	40	39	2.07	2.28	3.69
7	165	144	26	26	0.71	1.82	4.53
8	1009	1172	64	64	1.63	1.67	3.41
9	225	273	30	22	2.43	2.64	7.51
10	117	154	22	18	2.64	1.66	8.66
11	325	290	36	36	1.82	1.33	6.72
12	916	1109	61	61	1.61	3.15	10.02
13	224	208	30	30	1.27	3.26	17.05
14	89	66	19	18	2.00	2.45	3.29
15	326	313	36	35	2.11	2.58	7.96
16	326	157	36	35	1.44	2.66	14.09
17	354	339	38	33	2.31	2.94	10.82
18	66	117	16	16	0.61	2.62	13.71
19	152	227	25	22	2.05	1.35	15.68
20	320	351	36	36	0.81	2.18	6.15
<b>Average</b>	<b>389.8</b>	<b>407.5</b>	<b>36.7</b>	<b>35.3</b>	<b>1.61</b>	<b>2.31</b>	<b>8.30</b>

Table 5.7: Results of the proposed image registration method on 20 image pairs, each pair showing the same hemangioma at the same time.

matches are classified as inliers which corresponds to a percentage of  $\sim 96\%$ .

In Figure 5.4 for image pairs (1)-(3) the sensed image, the reference image and the difference image between the transformed sensed and the reference image are shown. It can be seen that the difference images show small disparities in the range of  $\sim 0.1$  inside and around the hemangioma region (note that larger local differences are caused by highlights on the hemangiomas). An exception is Figure 5.4(1) where larger differences up to a value of  $\sim 0.35$  arise at the hemangioma border because the assumption of hemangioma planarity is violated.

### 5.2.3 Accuracy of Registration on Follow-up Images

**Setup:** To assess the accuracy of the proposed method for consecutive images of the same hemangioma taken at different times we applied our algorithm on six different image series. Naturally, the hemangioma appearance changes from one examination date to another, i.e. the hemangioma regresses more and more during follow-up. Each image series consists of three to five images, resulting in a total of 18 registrations marked by two

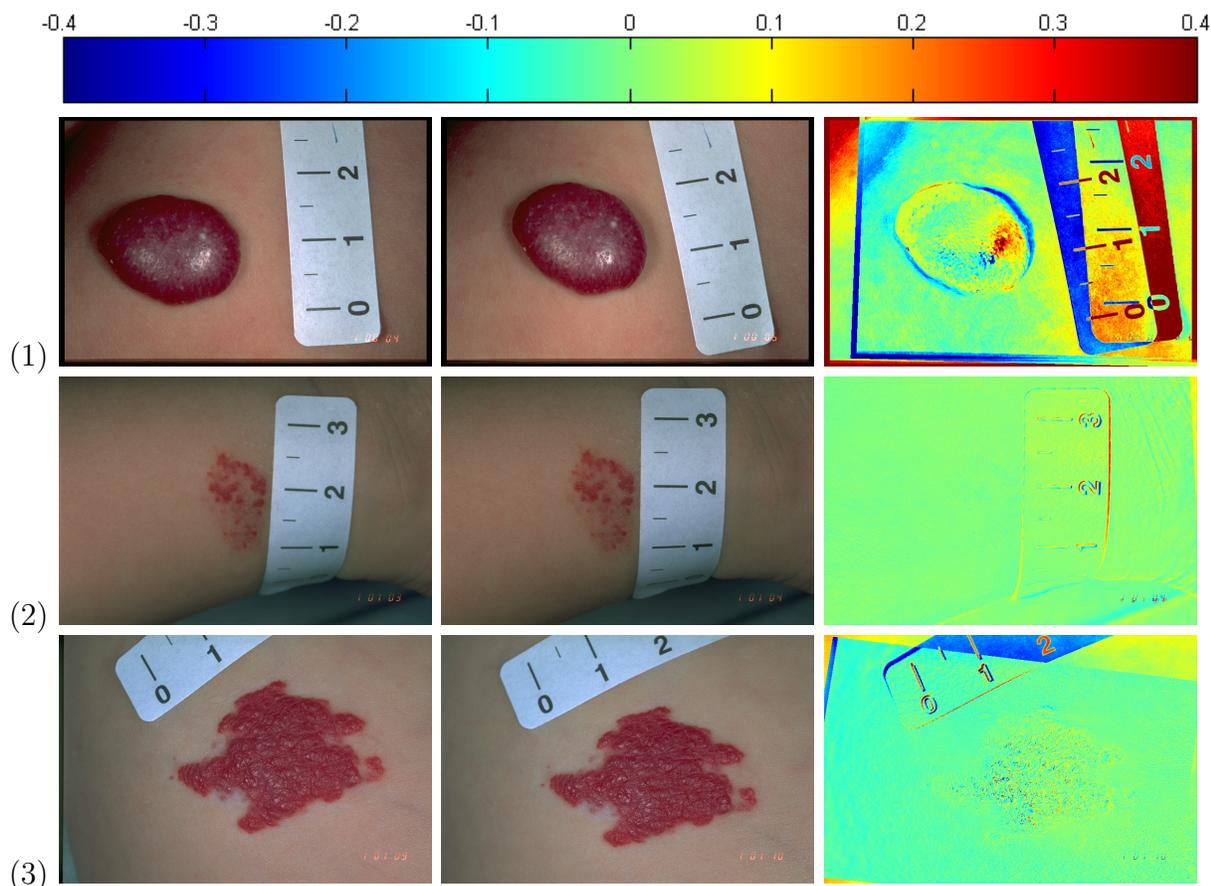


Figure 5.4: Difference images (1)-(3) of image pairs showing the same hemangioma at the same time. The colorbar is shown at the top.

characters, patient (1,2,3,4,5,6) and position in the sequence of registrations (A,B,C,D). For instance, 2C indicates the registration of the images from the third and the fourth examination of patient 2. The same error metrics as in Section 5.2.2, apart from border error, are measured and listed in Table 5.8. Additionally, a further error metric measuring the consistency of three circularly concatenated registrations is used, which is illustrated in Figure 5.5. In this experiment two images from examination 1 (Image  $A$  and Image  $A'$ ) and one from the subsequent examination 2 (Image  $B$ ) are used. The three homographies  $H_1$ ,  $H_2$  and  $H_3$  between the images are computed and the composite homography  $H = H_3H_2H_1$  is built. In the absence of error,  $H$  is the identity matrix and every point in Image  $A$  is not displaced by a transformation with it. Inevitably, there is an error which can be measured by the average displacement of points in Image  $A$ . Therefore, a set of reference points is equally placed in 10 pixel distances inside the hemangioma region of Image  $A$  (green points in Figure 5.5) and transformed by the composite homography  $H$ . The error is computed as the average Euclidean distance between reference points and their corresponding transformed points, listed as *Average Reference Points Displacement* in Table 5.8. This error is not stated for the registrations 3C, 4B and 4C because in these cases only one image from examination 1 is available.

Image	Interest Points Sensed Image	Interest Points Reference Image	Initial Matches	Inliers	Average Distance Error of Inliers (pixels)	Average Distance Error of 5 Test Points (pixels)	Average Reference Points Displacement (pixels)
1A	500	590	45	17	3.03	7.96	2.41
1B	590	671	49	28	2.79	3.96	8.67
1C	671	664	52	31	2.90	6.45	5.79
2A	1441	945	76	30	3.30	7.39	2.00
2B	945	1350	61	15	2.77	8.77	5.42
2C	1350	1497	73	35	2.41	12.94	5.67
2D	1497	187	77	7	1.41	354.84	440.38
3A	393	266	40	12	3.61	3.95	3.60
3B	266	191	33	18	2.57	4.92	3.76
3C	191	550	28	16	2.70	5.21	
4A	363	477	38	8	3.33	10.08	11.50
4B	477	213	44	11	3.97	4.69	
4C	213	767	29	6	0.94	196.43	
5A	320	509	36	21	2.51	3.44	2.70
5B	509	421	45	14	2.91	4.32	6.02
6A	169	375	26	7	2.63	8.37	7.50
6B	375	942	39	9	2.38	8.54	8.70
6C	942	320	61	8	1.79	5.50	6.82
<b>Average</b>	622.89	607.50	47.33	16.28	<b>2.66</b>	<b>36.54</b>	<b>34.73</b>
<b>Average (without 2D and 4C)</b>	593.88	623.810	46.63	17.50	<b>2.85</b>	<b>6.65</b>	<b>5.75</b>

Table 5.8: Results of the proposed image registration method on 6 image series with overall 18 registrations.

**Results:** Not surprisingly, the results are worse than for the image pairs tested in Section 5.2.2 since the content of the images changes from one time to another. Nevertheless, the average distance error of 5 test points and the average reference points displacement never exceeds 12 pixels ( $\sim 0.9$  mm), with the exception of the registrations 2D and 4C. In these two cases the hemangiomas have changed too much to obtain reliable matches and a meaningful homography. By excluding these two outliers we achieve an average reference points displacement of 5.75 pixels ( $\sim 0.4$  mm). The results of the match detection deteriorate, indicated by a lower fraction of initial matches finally classified as inliers ( $\sim 38$  %).

Figure 5.6 and 5.7 shows exemplarily for the follow-up images of patient 1 and 2 the sensed image, the reference image and the difference image between the transformed sensed and the reference image. Naturally, compared to the difference images of image pairs show-

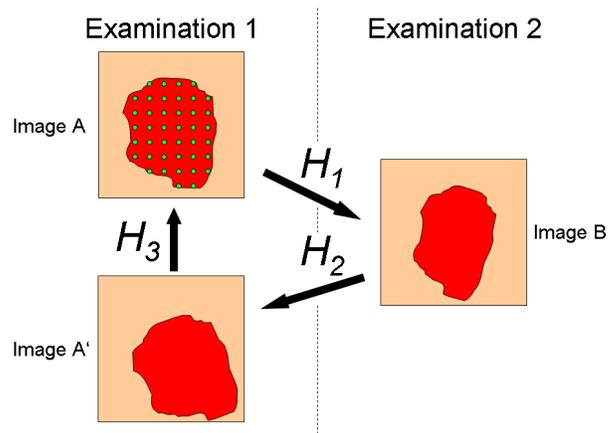


Figure 5.5: Illustration of the scheme for testing consistency of three circularly concatenated registrations.

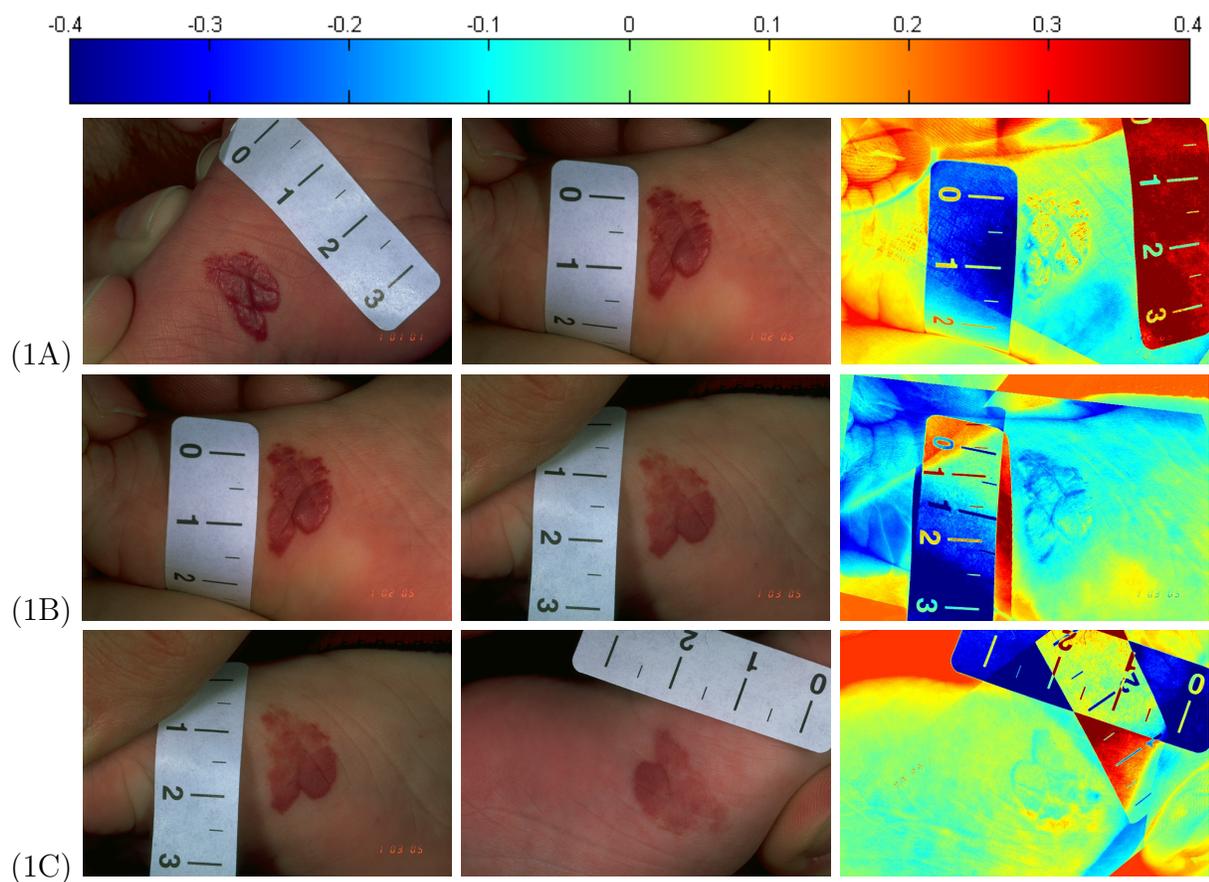


Figure 5.6: Difference images (1A)-(1C) of patient 1 in Table 5.8 showing the same hemangioma at different times. The colorbar is shown at the top.

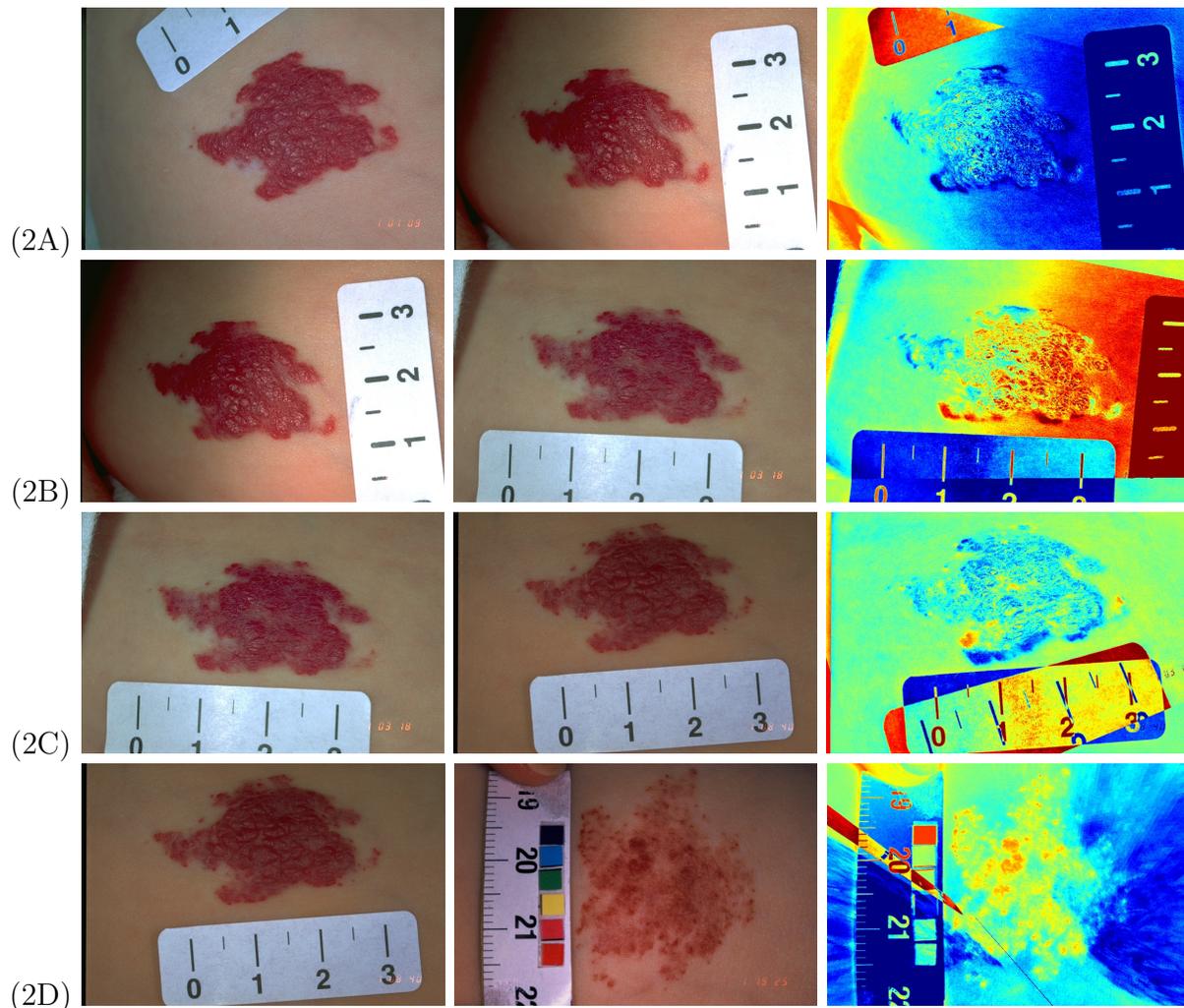


Figure 5.7: Difference images (2A)-(2D) of patient 2 in Table 5.8 showing the same hemangioma at different times.

ing the same hemangioma at the same time of Figure 5.4, higher differences occur since the appearance of a hemangioma changes from one time to another. For example, in Figure 5.6(1C) differences occur at the regressing regions of the reference image. In Figure 5.7(2C) the lower parts of the hemangioma are not correctly registered, indicated by higher differences in this border region compared to the other border regions of the difference image. The reason for such local errors is an irregular distribution of matches (and therewith inliers) in the hemangioma region, i.e. regions not represented sufficiently by matches are neglected by the homography estimation and tend to local misregistrations. Nevertheless, it must again be noted that the combination of SIFT, RANSAC and homography transformation model provide for a robust registration of follow-up hemangioma images with changing appearances. By the design of the algorithm, matching interest points of (mostly) unchanged hemangioma regions are implicitly detected and used for homography estimation.

## 5.3 Detection of Regressing Hemangioma Regions

This section contains several experiments for the evaluation of the proposed regression detection method. Initially, the method is applied on a set of synthetic follow-up images in Section 5.3.1. In Section 5.3.2 the method’s accuracy on real clinical images is tested by means of various error metrics. The suitability of including the proposed registration procedure in the detection scheme is shown in Section 5.3.3. Finally in Section 5.3.4 the method’s precision is evaluated on image pairs showing the same hemangioma at the same time.

### 5.3.1 Accuracy of Regression Detection on Synthetic Follow-up Images

**Setup:** As a proof of concept, the method proposed for regression detection was tested on three synthetic image series showing an increasing regression area over time. For each series a synthetic image of the type described in Section 5.1.2 was taken as initial image. Consecutive images were generated by a random 2-dimensional affine transformation (translation, rotation and scaling) of the hemangioma area to simulate viewpoint changes. The regression was simulated by painting a gray region inside the hemangioma area. Again gaussian noise with mean  $\mu = 0$  and variance  $\sigma^2 = 0.5$  was added to the images and they were smoothed with a  $7 \times 7$  gaussian filter with standard deviation  $\sigma = 1.0$ . Each series consists of four images and for each image of follow-up the percentage of the hemangioma area showing a graying was measured with the proposed method and compared to the known real values. The results are listed in Table 5.9. One of the three synthetic follow-up series is shown in Figure 5.8 where the automatically detected regressions are marked by a blue border.

**Results:** The average relative error on the three follow-up series is 2.73 %. Similar to the results achieved with the area measurement of synthetic images in Section 5.1.2, the

Image	Real Percentage	Measured Percentage	Error Percentage
1A	3.65	3.53	3.16
1B	14.55	14.87	2.19
1C	24.97	25.23	1.04
2A	11.86	11.35	4.35
2B	38.07	36.95	2.96
2C	55.03	53.50	2.79
3A	10.66	11.19	4.98
3B	14.72	15.12	2.75
3C	66.60	66.82	0.32
Average			<b>2.73</b>

Table 5.9: Results of regression detection on three follow-up series with synthetic images.

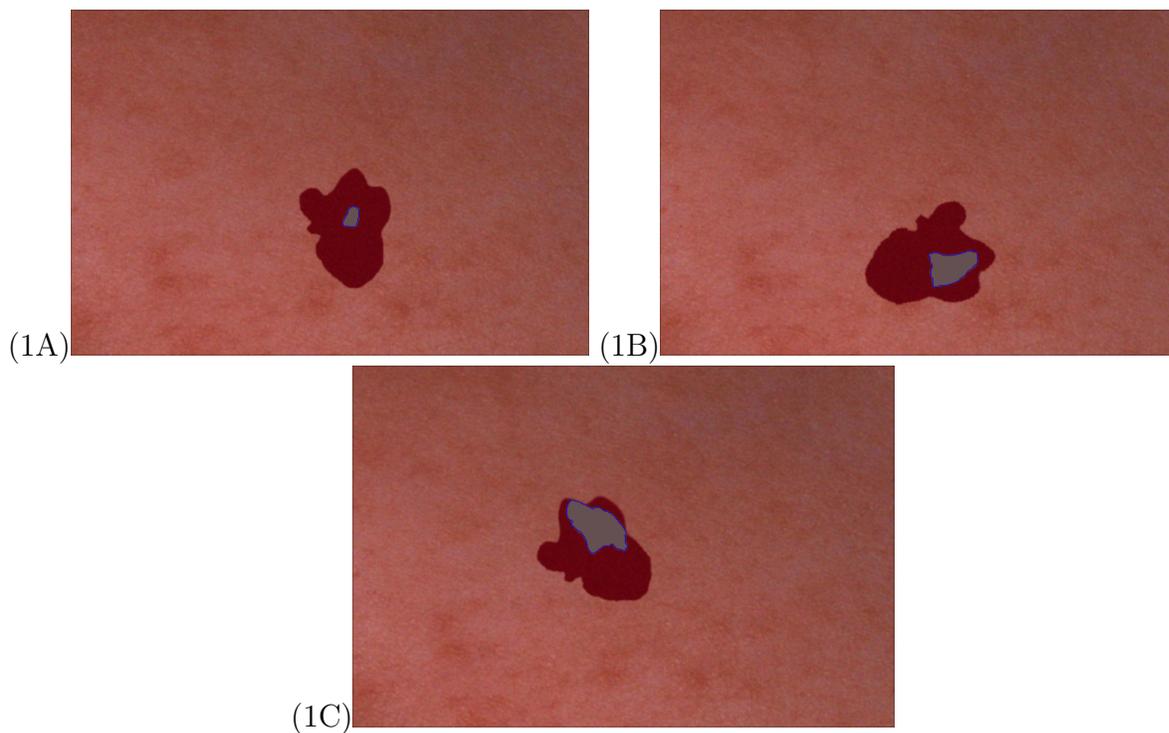


Figure 5.8: The second, third and fourth image of a synthetically generated follow-up series showing increasing regressions (gray regions). The area automatically detected by the proposed method is indicated by a blue border.

error is comparatively higher on images with smaller regressions, therefore in all three series the error consecutively decreases.

### 5.3.2 Accuracy of Regression Detection on Real Follow-up Images

**Setup:** For evaluation of the regression detection procedure the algorithm was tested on the follow-up images of Table 5.8. The follow-up images of patient 2 and 4 were excluded for of two reasons: (1) image quality impedes a consistent manual annotation of regressing regions and (2) these series show a misregistration between the last two images inhibiting an evaluation of the whole image series. Thus, in total 11 single images are evaluated in terms of the error metrics presented in Section 5.1.3. For an evaluation of the whole dataset a *leave-one-out cross-validation* scheme is applied: for every image of a series a new network is trained with the data from all series except the one containing the present image and tested on this image. The overall results are shown in Table 5.10. As described in Section 4.1, a neural network classifier is used. The network consists of 2 hidden layers with 20 and 10 hidden units, respectively. The network is trained by quasi-newton optimization [2] and training performance is measured by the mean squared error in combination with a weight decay performance ratio  $\gamma = 0.25$  to avoid overfitting (see Section 4.1.1).

Image	Error Rate (%)	False Pos. Rate (%)	False Neg. Rate (%)	Absolute Area Difference ( $cm^2$ )	Border Error (%)
1A	13.7	12.1	1.7	0.2508	39.9
1B	9.5	5.8	3.8	0.0576	20.7
1C	6.7	2.9	3.8	0.0310	16.3
3A	5.9	4.5	1.5	0.0712	20.0
3B	8.0	2.6	5.5	0.0835	60.2
3C	15.7	13.6	2.1	0.2403	38.6
5A	13.6	1.7	11.8	0.1996	38.6
5B	5.9	2.8	3.1	0.0071	17.4
6A	7.5	3.3	4.2	0.0315	15.9
6B	10.1	3.6	6.5	0.1126	17.6
6C	11.4	1.7	9.7	0.3204	18.2
<b>Average</b>	<b>9.8</b>	<b>5.0</b>	<b>4.9</b>	<b>0.1278</b>	<b>27.6</b>

Table 5.10: Results of the proposed method for regression detection on 4 image series with overall 11 registrations.

Since the overall goal of the proposed method is the automatic determination of the area/percentage of the hemangioma showing regression signs, these values are also reported for the same image series. They are listed separately in Table 5.11.

Image	Regressing Area ( $cm^2$ )	Total Area of Hemangioma ( $cm^2$ )	Percentage
1A	1.08	1.08	100.00
1B	1.38	1.38	100.00
1C	1.36	1.36	100.00
3A	0.78	1.03	75.34
3B	0.28	1.06	26.82
3C	1.09	1.11	97.84
5A	0.50	0.65	77.10
5B	0.77	0.77	99.76
6A	1.47	1.66	88.42
6B	2.11	2.20	96.10
6C	2.24	2.31	97.12

Table 5.11: Absolute and relative determined areas of regressing hemangioma regions of the images from Table 5.10.

**Results:** The method shows an average border error of 27.6 % which is in the range of the border error of the hemangioma segmentation method (32.1 %, see Table 5.3). The average absolute error made in the surveying of the regression region lies at  $0.1278 cm^2$ , where the average size of regression areas in the given data is  $1.2015 cm^2$ .

Figure 5.9 shows the results obtained on the single images. Here the region of manual

detection is marked by a green border whereas the automatically detected region is marked by a blue border. The largest error occurs in image (3B) with a border error of 60.2 %, but this image exhibits a strong oversaturation caused by the inadequate image acquisition process.

Regarding the test reporting the area/percentage of the regressing region, it can be seen that the fraction of the hemangioma showing a regression increases or remains constant in all follow-up series. The only exception is the previously mentioned poor-quality image (3B). For patient 1 the algorithm assesses already the whole hemangioma area of the first image (1A) as regressing (percentage = 100 %). It must be stated that during the usual five-week-interval between examinations the child grows and therewith also the hemangioma on the skin. For that reason at the examinations the child's current body height is measured to normalize the data with it. This was not taken into account for these experiments.

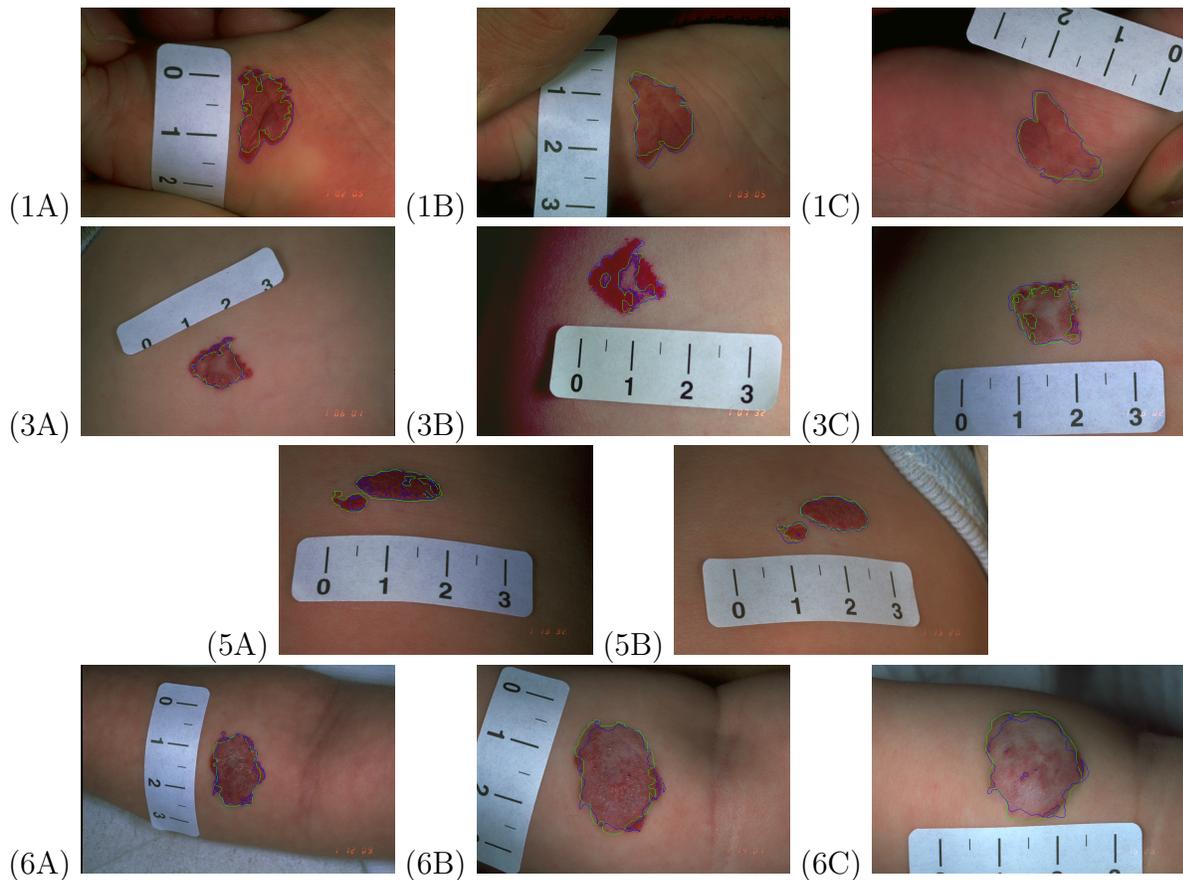


Figure 5.9: Results of the automatic regression detection (green: manual detection, blue: automatic detection).

### 5.3.3 Accuracy of Regression Detection on Follow-up Images without Registration

**Setup:** As discussed in Section 4.1.2, the change of hemangioma color values is a discriminative feature for the classifier. To verify this, the experiment reported in Section 5.3.2 was repeated without using difference features, i.e. only  $G$ ,  $a^*$  and  $abdist$  were used.

Error Rate	False Pos. Rate	False Neg. Rate	Absolute Area Difference	Border Error
16.3%	6.9%	9.4%	0.3486 $cm^2$	50.1%

Table 5.12: Average results of the proposed method for regression detection on 4 image series with overall 11 registrations (without difference features).

**Results:** The average errors are listed in Table 5.12. It can be seen, that all error values deteriorate compared to the values obtained by using registration and difference features. Border error increases from 27.1 % to 50.1 % and absolute area difference from 0.1278  $cm^2$  to 0.3486  $cm^2$ . This is a strong indication that the registration and the thereby available difference features taking follow-up examinations into account are a necessary prerequisite for a reliable regression detection.

### 5.3.4 Precision of Regression Detection on Hemangioma Images taken at the same Time

**Setup:** Since the manual determination of regressing regions is an uncertain and difficult task (even for dermatologists), this experiment deals with the comparison of surveyed regression regions in image pairs showing the same hemangioma, consistent with the experiment reported in Section 5.1.4. Again the absolute difference and variation coefficient between the two measurements are calculated. (3B) was excluded from this test because no second image was available for that case.

**Results:** All values are listed in Table 5.13. Compared to the hemangioma segmentation results, the variation coefficient is even less (7.40 % to 8.82 % in Table 5.5). Also the absolute difference shows a comparable value (0.1134  $cm^2$  to 0.0775  $cm^2$ ). Although no manual reference measurements exist for this test, the error is considered to be caused mainly by variations in the image acquisition procedure, likewise the results for the precision of the hemangioma segmentation.

## 5.4 Discussion

As part of the hemangioma assessment procedure, the first task is to determine and measure the whole skin region affected (Section 5.1). By the design of the algorithm, errors in area computation arise due to three different reasons:

Image Series	Absolute Area Difference ( $cm^2$ )	Variation Coefficient (%)
1A	0.0906	6.18
1B	0.0854	4.56
1C	0.2570	14.75
3A	0.1165	9.19
3C	0.0986	6.63
5A	0.0467	7.83
5B	0.1272	13.02
6A	0.1597	7.14
6B	0.0120	0.39
6C	0.1402	4.33
Average	<b>0.1134</b>	<b>7.40</b>

Table 5.13: Differences in automatically determined areas of regression between two images of the same hemangioma.

1. **Error in image scale computation:** Tests in Section 5.1.1 have shown that scale computation is accurate (average error rate of 1.55 %), with very few exceptions where the ruler was inadequately placed on the skin or photographed. Therefore, this type of error is assumed to have the lowest impact on area measurement.
2. **Error in hemangioma segmentation:** Hemangioma segmentation works reliably and accurately on saturated hemangiomas with no or few regressions, as shown in Section 5.1.3 where 54 out of 90 test images could be segmented with a border error of less than 20 %. On the other side, the algorithm can fail completely on regressing hemangiomas because of the lower contrast between skin and hemangioma regions. However, typically the first images during follow-up studies exhibit no or few regressions and can therefore be segmented accurately. The affected skin area in later examinations can then be determined by the proposed registration method.
3. **Variation in image acquisition:** For an optimal measurement both ruler and hemangioma have to be situated parallel to the view plane of the camera. Whereas this is naturally violated by the fact that hemangioma and skin are never perfectly planar, errors are also caused by an inadequate positioning of the camera. The tests made on the image pairs in Section 5.1.4 show that the variation coefficient achieved with the presented method (8.82 %) is not far away from the one achieved with a manual surveying (4.53 %). An error of  $\sim 4\%$  can be seen as the natural limitation of the method, i.e. the minimal achievable error rate.

**Registration of follow-up images:** Experimental results reported in Section 5.2 show that the majority of the images can be registered without considerable errors. The average reference point displacement of  $\sim 0.4mm$  accumulated by three subsequent registrations is acceptable with an average hemangioma diameter of  $\sim 14mm$ . Sources of error are (1) the necessary assumption of hemangioma planarity due to the homography transformation model and (2) global or local misregistrations caused by strong changes in hemangioma

appearance and a thereby high fraction of false matches, either only in specific regions or on the whole hemangioma. However, the proposed method is designed to deal with changing hemangioma appearances and the tests show that it is applicable on the given follow-up data, although robustness of the method decreases in the end of follow-up series when intervals between examinations go up to 6 months. Here the strong changes in hemangioma appearance can result in a failure of the registration (see for example Figure 5.7(2D)).

**Regression detection:** The tests on the detection of regressing regions (Section 5.3) also verify that the registration is accurate enough to support the neural network in classification between regressing and not regressing regions by means of difference features. With difference features the classifier’s performance is improved, observable by a border error reduction from 50.1 % to 27.6 %. Nevertheless, an important aspect of regression detection is the uncertainty of the manual annotation (the so-called ground truth) and therefore the use of the border error as performance measure. However, by analyzing the determined absolute areas of the detected regression regions in follow-up series (see Table 5.11), an enlargement of regressing regions is reported. The uncertainty of ground truth annotation is overcome by a comparison of determined regression areas, measured on image pairs showing the same hemangioma at the same time. Because of the similar error values (variation coefficient of 7.40 % compared to 8.82 % of the hemangioma segmentation method), it is assumed that this error is again mainly caused by variations in the image acquisition procedure, i.e. different lighting conditions and camera viewpoints, although no manual reference measurements were available.

**Comparison with other methods:** As stated earlier, the image-based automatic assessment of cutaneous hemangiomas and the proposed scheme of detecting and measuring regressions by the use of registered follow-up images was not covered before. Thus, no direct comparisons to other methods can be made. However, other skin lesion types were addressed in the past and there are some comparable results despite different input data, experimental setups and algorithm designs: in [44] border error was used as performance metric for melanoma segmentation and the authors state that a border error of less than 50% is considered as sufficient by dermatologists. They report a border error of less than 50 % in 55 out of 66 images. In our results for hemangioma segmentation (Table 5.3) 75 out of 90 segmentations have a border error of less than 50 % which corresponds to the same percentage of 83 %. Regarding the area measurement of skin lesions, in [34] a method for psoriasis surveying was developed showing an error of less than 4 % to reference measurements in all cases. Although an exact comparison cannot be made because of nonexistent reference measurements, it must be noted that in the surveying of hemangioma/regression regions we obtain higher variation coefficients of 8.8 and 7.4 %, respectively. However, it is not clear to which extent their method is influenced by the image acquisition.

Generally, the method can be expected to outperform manual measurements during clinical trials. According to dermatologists, during an examination the hemangioma area is

simply determined by measuring the hemangioma's height and width with a tape measure. The percentage of the hemangioma area showing a graying or desaturation (i.e. the regression regions) is then estimated by the attending physician. Obviously, this kind of area measurement using a tape measure is less accurate than an automatic measurement on digital images. The determination of the regression regions percentage as part of the healing process assessment can be made more consistent as well, since estimations of several different dermatologists impede an objective evaluation during clinical trials. Nevertheless, in the future a detailed medical evaluation of the results achieved has to be done by dermatologists. Moreover, repeated measurements by dermatologists are necessary to quantify the improvement of the presented method.

# Chapter 6

## Conclusion

In this thesis a new system for the automatic image-based assessment of cutaneous hemangiomas is presented. Cutaneous hemangiomas are a type of skin disease that causes an increased growth of blood vessels beneath the skin in local regions. The procedure is able to measure the current area of the affected skin and to detect and survey regions showing a regression during follow-up. The area measurement is based on the image scale computation by means of a ruler attached to the skin and visible in the image to be examined. Both the detected hemangioma and the regression regions can then be surveyed by a determination of the respective image regions. The accurate measurement of the affected skin region, and in particular the development of the hemangioma during the follow-up examinations is crucial during clinical trials and to support an effective treatment.

The motivation for the proposed method was to make the assessment of hemangioma development more accurate and objective. Thereby the quality of both clinical trials and long-term studies can be improved by a more consistent evaluation of the effect of therapies. The current clinical standard for the assessment of the hemangioma's healing process is the simple manual measurement of its height and width and an estimation of the percentage showing a regression by the treating physician. This area estimate does not account for the irregular shape, which becomes more relevant during regression, and the percentage estimation is based on a rough visual estimate. The method proposed in this thesis measures the actual shape and area of the hemangioma and therefore provides a more accurate measurement of its size, which is not affected by inter- or intra reader variation, and can quantify even small local structure of the hemangioma. The detection of regressions is made more consistent by decreasing the subjectivity of a manual regression quantification caused by differences of opinion between observers.

The segmentation of the hemangioma region is accomplished by a pixel-wise classification scheme, using the pixel's color values as input for a single-layer perceptron. Likewise, for the determination of regions showing a regression a pixel-wise classification is also applied. Since differences between regressed and unregressed regions are less distinct than the differences between hemangioma regions and healthy skin, for this task a more sophisticated classifier, a neural network (multi-layer perceptron), is used. Moreover, the classification is supported by a comparison to the first image of the follow-up series by means of difference features. Difference features are obtained by image subtraction of the examined image with the first follow-up image, aligned to each other by robust image registration of

the hemangioma area. The proposed method for image registration has to cope with the changing hemangioma appearances between examinations. This is achieved by a feature-based method using SIFT features to obtain reliable matches between consecutive images. The inevitable occurrence of incorrect matches caused by the varying image content is mastered by the use of a RANSAC scheme for robust homography estimation describing the transformation between the two images.

Although no manual reference measurements for an exact quantification of the achieved improvement were available, the experimental results indicate the applicability on clinical trials. 60 % of the test images could be segmented with a border error of less than 20 % and the algorithm shows high robustness especially on the first images of follow-up, an important property for the proposed workflow where the region of interest (i.e. the hemangioma region) is determined in the first image and transformed via image registration for an assessment of the subsequent images. The registration scheme itself shows high accuracy with an error rate of  $\sim 0.4mm$  on hemangiomas with an average diameter of  $\sim 14mm$ . For the detection of regressions only a limited set of 4 follow-up series were available. Nevertheless, the obtained border error of 27.6 % and variation coefficient of 7.4 % are comparable to the results of the hemangioma segmentation (32.1 % and 8.82 %, respectively).

In general, the individual methods as well as the whole assessment procedure suffer from the partial low image quality since the image acquisition process during examinations is usually less than perfect from a technical point of view: image quality is degraded by imperfect illumination conditions causing specular reflections (highlights), an inconsistent illumination of the hemangioma and the inappropriate photo development of digitalizing analog images with a scanner. Furthermore, an inadequate choice of camera viewpoint is a significant source of inaccuracy since for area measurement both the hemangioma and the ruler are assumed to be parallel to the view plane of the camera.

Regarding the usage of the proposed method in clinical practice, it must be stated that despite the reported robustness, visual human control of the results is necessary. However, by the design of the procedure, corrections of false results can easily be made by the physician: since the hemangioma segmentation and regression detection are based on classification, the output of the respective classifier can be interpreted as the confidence in the chosen class. Therefore, a correction of the segmentation can be achieved by a simple adjustment of the threshold dividing the two classes. As a further aspect, parts of the procedure can be replaced by manual methods. For instance, an incorrect registration can be compensated by a manual definition of corresponding points in the images.

Although experiments in this thesis were only conducted on follow-up images of cutaneous hemangiomas, other kinds of lesions can be examined as well. The classification- and registration scheme makes the method highly adaptable to other lesion types, provided that an adequate amount of annotated training data is available. The performance on cutaneous hemangiomas could be improved as well with a larger amount of follow-up series.

Future research should focus on a more extensive quantitative comparison of the proposed method with the manual state of the art assessment, to prepare the application for clinical practice. Both clinical trials and patients can be expected to benefit from the improved accuracy and the higher specificity, allowing for the quantification of fine changes during

the course of the disease.

# Bibliography

- [1] E. Barnard. Optimization for training neural nets. *IEEE Transactions on Neural Networks*, 3(2):232–240, 1992.
- [2] C. M. Bishop. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.
- [3] J. Canny. A computational approach to edge detection. *PAMI*, 8(6):679–698, 1986.
- [4] N. Cascinelli, M. Ferrario, T. Tonelli, and E. Leo. A possible new tool for clinical diagnosis of melanoma: The computer. *J. Amer. Acad. Dermatol.*, 16(2):361–367, 1987.
- [5] M.E. Celebi, H.A. Kingravi, B. Uddin, H. Iyatomi, Y.A. Aslandogan, W.V. Stoecker, and R.H. Moss. A methodological approach to the classification of dermoscopy images. *Computerized Medical Imaging and Graphics*, 31:362–373, 2007.
- [6] K.G. Chiller, D. Passaro, and I.J. Frieden. Hemangiomas of infancy: clinical characteristics, morphologic subtypes, and their relationship to race, ethnicity, and sex. *Eur Arch Otorhinolaryngol*, 258:141–149, 2001.
- [7] T.M. Cover and P.E. Hart. Nearest neighbor pattern classification. *IT*, 13(1):21–27, January 1967.
- [8] G. R. Day and R. H. Barbour. Automated melanoma diagnosis: Where are we at? *Skin Res. Technol.*, 6:1–5, 2000.
- [9] D. Delgado, B. K. Ersbøll, and J. M. Carstensen. S.H.A.R.P: A smart hierarchical algorithm to register psoriasis. In *Proc. of Int. Workshop on Systems, Signals and Image Processing*, pages 43–46, 2004.
- [10] B.A. Drolet, N.B. Esterly, and I.J. Frieden. Hemangiomas in children. *New England Journal of Medicine*, 341:173–181, 1999.
- [11] M.A. Fischler and R.C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
- [12] N. I. Fisher. *Statistical Analysis of Circular Data*. Cambridge University Press, 1993.

- [13] H. Ganster, P. Pinz, R. Rohrer, E. Wildling, M. Binder, and H. Kittler. Automated melanoma recognition. *IEEE Transactions on Medical Imaging*, 20(3):233–239, March 2001.
- [14] D.D. Gomez, C. Butakoff, B. Ersbøll, and J.M. Carstensen. Automatic change detection and quantification of dermatological diseases with an application to psoriasis images. *Pattern Recognition Letters*, 28(9):1012–1018, 2007.
- [15] R. C. Gonzalez, R. E. Woods, and S. L. Eddins. *Digital Image processing using MATLAB*. Pearson Prentice Hall, New Jersey, 2004.
- [16] A. Green, N. Martin, J. Pfitzner, M. O’Rourke, and N. Knight. Computer image analysis in the diagnosis of melanoma. *J. of the American Academy of Dermatology*, 31(6):958–964, December 1994.
- [17] W. Guo and Y.A. Aslandogan. Mining skin lesion images with spatial data mining methods. Technical Report CSE-2003-19, Department of Computer Science and Engineering, University of Texas at Arlington, July 2003.
- [18] T. P. Habif. *Clinical Dermatology*. The C.V. Mosby Company, 2nd edition, 1990.
- [19] G.A. Hance, S.E. Umbaugh, R.H. Moss, and W.V. Stoecker. Unsupervised color image segmentation with application to skin tumor borders. *IEEE Engineering in Medicine and Biology*, 15(1):104–111, January/February 1996.
- [20] R.I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, 2nd edition, 2004.
- [21] S.L. Horowitz and T. Pavlidis. Picture segmentation by a tree traversal algorithm. *Journal of the ACM*, 23(2):368–388, April 1976.
- [22] T.D. Jones and P. Plassmann. An active contour model for measuring the area of leg ulcers. *IEEE TMI*, 19(12):1202–1210, 2000.
- [23] J.M. Kasson and W. Plouffe. An analysis of selected computer interchange color spaces. *ACM Transactions on Graphics*, 11(4):373–405, October 1992.
- [24] P. Krizek, J.V. Kittler, and V. Hlavac. Feature selection based on the training set manipulation. In *Proc. of ICPR 2006*, volume 2, pages II: 658–661, 2006.
- [25] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.
- [26] I. Maglogiannis. Automated segmentation and registration of dermatological images. *Journal of Mathematical Modelling and Algorithms*, 2(3):277–294, 2003.
- [27] H. Maier and R. Neumann. Treatment of strawberry marks with flashlamp-pumped pulsed dye laser in infancy. *Lancet*, 347:131–132, 1996.

- [28] G. Maletti and B. Ersbøll. Change detection in registered psoriasis lesion image patterns. Technical report, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 2003.
- [29] G. Maletti and B. Ersbøll. Texture alteration detection in bitemporal images of lesions with psoriasis. Technical report, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 2003.
- [30] G. Maletti, B. K. Ersbøll, and K. Conradsen. A combined alignment and registration scheme of lesions with psoriasis. *Information Sciences*, 175(3):141–159, 2005.
- [31] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. In *Proc. of Conference on Computer Vision and Pattern Recognition*, pages 257–263. IEEE Computer Society, 2003.
- [32] F. Nachbar, W. Stolz, T. Merkle, A.B. Cognetta, T. Vogt, M. Landthaler, P. Bilek, O. Braun-Falco, and G. Plewig. The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions. *J. Amer. Acad. Dermatol.*, 30(4):551–559, 1994.
- [33] S.A. Pavlopoulos. New hybrid stochastic-deterministic technique for fast registration of dermatological images. *Medical & Biological Engineering & Computing*, 42(6):777–786, 2004.
- [34] J. Roening, R. Jaques, and J. Kontinen. Area assessment of psoriatic lesions based on variable thresholding and subimage classification. In *Proc. of Vision Interface '99*, 1999.
- [35] F. Rosenblatt. The perceptron: A perceiving and recognizing automaton. Report 85-460-1, Project PARA, Cornell Aeronautical Laboratory, Ithaca, New York, January 1957.
- [36] A.J. Round, A.W.G. Duller, and P.J. Fish. Colour segmentation for lesion classification. In *IEEE Proceedings of the Engineering in Medicine and Biology Society*, volume 2, pages 582–585, October/November 1997.
- [37] R.E. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. *Machine Learning*, 37(3):297–336, 1999.
- [38] P. Schmid and S. Fischer. Colour segmentation for the analysis of pigmented skin lesions. In *Proceedings of the Sixth International Conference on Image Processing and its Applications*, volume 2, pages 688–692, July 1997.
- [39] P. C. M. van de Kerkhof. On the limitations of the psoriasis area and severity index (PASI). *British Journal of Dermatology*, 126:205, 1991.
- [40] J.A. Werner, A.A. Duenne, B.J. Folz, R. Rochels, S. Bien, A. Ramaswamy, and B.M. Lippert. Current concepts in the classification, diagnosis and treatment of hemangiomas and vascular malformations of the head and neck. *Arch. Dermatol.*, 138(12):1567–1576, 2002.

- [41] L. Xu, M. Jackowski, A. Goshtasby, D. Roseman, S. Bines, C. Yu, A. Dhawan, and A. Huntley. Segmentation of skin cancer images. *Image and Vision Computing*, 17(1):65–74, January 1999.
- [42] S. Zambanini, G. Langs, R. Sablatnig, P. Donath, and H. Maier. Automatic surveying of cutaneous hemangiomas. In *Proc. of ICPR 2006*, volume 1, pages 1022–1025, Hong Kong, 2006.
- [43] S. Zambanini, G. Langs, R. Sablatnig, and H. Maier. Automatic robust registration of cutaneous hemangiomas for follow-up examinations. In *Proc. of 31st AAPR/OAGM Workshop*, volume 224, pages 121–128, Krumbach, Austria, 2007.
- [44] Z. Zhang, W.V. Stoecker, and R.H. Moss. Border detection on digitized skin tumor images. *IEEE Transactions on Medical Imaging*, 19(11):1128–1143, November 2000.
- [45] B. Zitova and J. Flusser. Image registration methods: a survey. *IVC*, 21(11):977–1000, 2003.