

Kernel Fisher Discriminant and Elliptic Shape Model for Automatic Measurement of Allergic Reactions

Heikki Huttunen¹, Jari-Pekka Ryyänen¹, Heikki Forsvik¹,
Ville Voipio¹, and Hisakazu Kikuchi²

¹ Department of Signal Processing, Tampere University of Technology, Finland

² Graduate School of Science and Technology, Niigata University, Japan

Abstract. A semiautomatic segmentation method for images of allergic reactions in skin prick test is proposed. The method is based on elliptic model for the shape of the wheal, and it uses the kernel Fisher discriminant for grayscale projection and for measuring the separability of the object and the background areas. Experiments indicate that the method is robust and the results are close to those obtained manually.

Keywords: Allergy Test, Kernel Fisher Discriminant, Elliptic Shape Model.

1 Introduction

The skin prick test is a standard allergy diagnosis method. Compared to its alternatives, e.g., measuring the antibody levels in the blood, it is simpler and the results are available immediately. In a typical skin prick test, multiple allergens are tested at the same time by placing drops of allergen extracts on the patient's skin (usually on the forearm) and piercing the skin with a small metal lancet. A positive reaction induces a raised itchy area, a *wheal*, whose size is used for estimating the sensitivity to certain allergen. Traditionally, a medical doctor uses a ruler to manually measure the size of the wheal. The test procedure assumes an elliptic shape, with possible elongated branches (called *pseudopodia*) disregarded and the result of the measurement is the mean of the major and minor axes of the imaginary ellipse [10]. An illustration of the measurement is in Figure 1. As can be seen, most practitioners apply pressure to improve the visibility.

Problems with the traditional method are that the measurement result is subjective and only one measurement can be taken at a single time instant (so all wheals are not measured simultaneously). The wheal growth speed may also have some clinical importance, which can not be exploited today. Additionally, no documentation is left from the reaction (besides the possibly biased measurements). Thus, automation of allergy test using image analysis is of interest.

Several approaches using digital photography and subsequent image analysis for skin erythema detection and melanoma detection have been proposed. There exists a lot of literature on melanoma segmentation [2,4,5], but only a few studies



Fig. 1. Manual allergy measurement

of measurement of allergic reactions from 2D pictures [8,9], 3D profiles [10] or other specialized imaging hardware [13]. Our concern is in low-cost 2D digital color photography. Among these, Roullot *et al.* [9] considers seven well known color spaces and compares the separability of the reaction from the background using a training database. As a result, they discover that the optimal dimension among the color spaces is the a^* -component of the $L^*a^*b^*$ color space. Using the extracted a^* -component, they use simple thresholding for segmenting the wheal.

Nischik *et al.* [8] also discover the $L^*a^*b^*$ color space most suitable for the wheal segmentation and use the standard deviations of the L^* and a^* components as the features for classification. The classifier is trained to separate between foreground (the wheal) and the background (healthy skin) using manually generated training data. The classifier output determines directly the boundary between the two regions.

Recent work by Celebi *et al.* consider finding optimal color transformation for extracting the foreground [2]. Although the paper concentrates on melanoma segmentation, the principle is applicable for other skin diseases, as well. The paper searches for optimal linear combination of the RGB-components, such that the output maximizes the separability of the foreground and background. The foreground and background are determined in each iteration using Otsu thresholding. Thus, the algorithm iterates all projections defined on a finite grid, and tests their performance by measuring the Fisher ratio of the foreground and background (which are determined using Otsu thresholding).

The method of [2] is an unsupervised method, which attempts to find the best projection without any user manual assistance. However, in our work we study the case, in which the user points the approximate center of the wheal. From the practical point of view this is acceptable, because it requires less work than the manual measurement. However, we plan to automate the detection of the wheal location in the future. Note also that in the temporal direction, clicking



Fig. 2. Proposed wheal detection framework

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

The method proposed in this paper consists of the steps in the block diagram of Figure 2 and are described in the paper as follows. In Section 2, the optimal grayscale projection is stated in terms of the Fisher Discriminant (FD). Section 2 also considers the Kernel Fisher Discriminant (KFD), which generalizes the FD using the kernel trick. After describing the optimal grayscale mapping, we define an elliptical shape model for the wheal in Section 3. Finally, in Section 4 we present experimental results and conclude with discussion in Section 5.

2 Grayscale Projection

The key problem when searching the borders of the wheal is the poor contrast between the wheal and skin. An example of a wheal is illustrated in Figure 3 (a). Although the wheal borders are barely visible, the shape becomes highlighted when mapped into grayscale in a suitable manner. Well known mappings for skin color processing include the hue component of the HSV color space (Figure 3 (b)) and the a^* component of the $L^*a^*b^*$ color space (Figure 3 (c)). In all projections, we have smoothed the RGB image by convolution with a disc shaped window of radius 5. However, these are more or less arbitrary, and variability in skin color and allergic reaction strength may decrease their applicability. Instead, training based projections may improve the separation further, and make it more invariant for all patients. An unsupervised method for finding a well-separating projection in terms of the Fisher criterion was proposed by Celebi *et al.* [2], whose result is shown in Figure 3 (d). In this case the coefficients are 1, -0.1 and -0.3 for red, green and blue channels, respectively.

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

The training set is acquired as follows. When the user has pointed the approximate location of the center of the wheal, a set of RGB values is obtained from the neighborhood. In our experiments, the training set of the wheal (\mathcal{S}_1) is obtained inside the circular neighborhood with the radius of 10 pixels. The training set of the healthy skin (\mathcal{S}_0) is acquired from pixels that are far away from the center. In our experiments this is done by taking all pixels located at a radius between 45 and 50 pixels from the center, as illustrated in Figure 3 (h).

The natural tool for optimally projecting the three-dimensions to grayscale is the *Fisher Discriminant*, [3]. Fisher discriminant finds the projection dimension

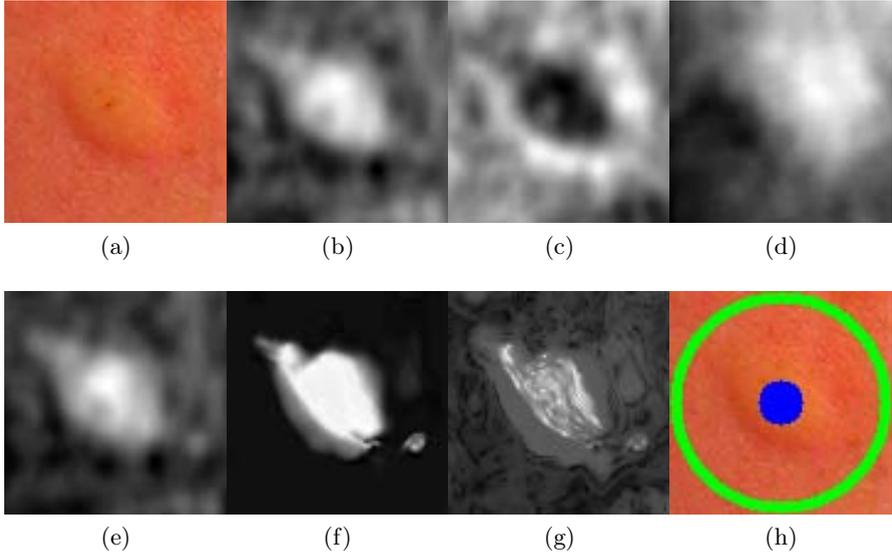


Fig. 3. The projection of the wheal in RGB color space. The original RGB image is shown in Figure (a), and (b) shows the hue component, (c) the a^* component of the $L^*a^*b^*$ color space and (d) is the projection proposed in [2]. Optimal Fisher discriminant projection is shown in Figure (e), and the results of its kernelized version are shown in figures (f) and (g) using RBF kernel with bandwidth σ selected using Silverman rule of thumb and with fixed $\sigma = 0.5$. The training sets are obtained from areas shown in (h), where the blue center is the foreground sample region and the green circle is the background sample region.

\mathbf{w} that maximizes the separability of the classes in terms of the ratio of the between-class-variance and within-class-variance; i.e., the so called Fisher ratio:

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}, \quad (1)$$

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

$$\mathbf{w} = \mathbf{S}_W^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0), \quad (2)$$

where $\boldsymbol{\mu}_1 \in \mathbf{R}^3$ and $\boldsymbol{\mu}_0 \in \mathbf{R}^3$ are the sample means of \mathcal{S}_1 and \mathcal{S}_0 . An example of the result of the Fisher discriminant projection is shown in Figure 3 (e).

The FD is a special case of so called *Kernel Fisher Discriminant* (KFD) [7,11], which is a kernelized version of the standard FD. As all kernel methods, the KFD implicitly maps the original data into a high-dimensional feature space and finds the optimally separating manifold there. Using the implicit mapping via the *kernel trick*, the explicit mapping can be avoided, which allows calculating the FD even in an infinite-dimensional space.

In practice the KFD can be calculated implicitly by substituting all dot products with a kernel function $\kappa(\cdot, \cdot)$. It can be shown, that all positive definite kernel functions correspond to a dot product after transforming the data to a feature space \mathcal{H} with mapping $\Phi(\cdot)$ [11]. The feature space \mathcal{H} can be very high dimensional, and the use of the projection vector \mathbf{w} directly may be impractical or impossible. Instead, the famous *Representer theorem* guarantees that the solution can be represented as a linear combination of the mapped samples [11]. Thus, the Fisher ratio in the feature space is based on the weights of the samples α instead of the weights of the dimensions:

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

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

Similar solution as the one for the Fisher discriminant in Eq. (2) can be found also for this case, [11]. However the inversion becomes more difficult, since the dimension of the weight vector α is now the number of the collected training samples. Therefore, we need a regularization term $\lambda \mathbf{I}$, where λ is a small positive scalar and \mathbf{I} is the $N \times N$ identity matrix. In our notation this yields the solution

$$\alpha = (\mathbf{Q}^T \mathbf{S}_W^\Phi \mathbf{Q} + \lambda \mathbf{I})^{-1} \mathbf{Q}^T (\mu_1^\Phi - \mu_0^\Phi), \tag{4}$$

where $\mu_1^\Phi \in \mathcal{H}$ and $\mu_0^\Phi \in \mathcal{H}$ are the sample means of the mapped wheal and skin samples, respectively. It is straightforward to show, that Eq. (4) can be expressed in terms of dot products and thus the kernel trick, [7,11]. Also the actual projection of a test sample $\mathbf{x} \in \mathbf{R}^3$ can be expressed through the kernel as $y = \alpha^T \mathbf{Q}^T \Phi(\mathbf{x}) = \sum_{i=1}^N \alpha_i \kappa(\mathbf{x}_i, \mathbf{x})$.

There are various alternatives for the kernel function $\kappa(\cdot, \cdot)$, among which the most widely used are the polynomial kernels and the Radial Basis Function (RBF) kernel. We experimented with various kernels, and found out that the polynomial kernels do not increase the separation significantly when compared with the linear kernel, which is equivalent to the traditional FD. In other words, all low-order polynomial kernels produce a projection very similar to the first order kernel, shown in Figure 3 (e). However, the separation seems to improve with the RBF kernel

$$\kappa(\mathbf{u}, \mathbf{v}) = \exp\left(-\frac{\|\mathbf{u} - \mathbf{v}\|^2}{2\sigma^2}\right). \tag{5}$$

There are two parameters in the KFD projection with RBF kernel: The regularization parameter λ and the kernel width σ^2 . Since there exists a lot of training data in our case, it seems to be less sensitive to the regularization parameter λ than the width σ^2 . In our experiments we set the value of $\lambda = 10^{-5}$, and if the condition number of the matrix in Eq. (4) indicates that the matrix is close to singular, the value of λ is increased ten-fold until the inversion succeeds.

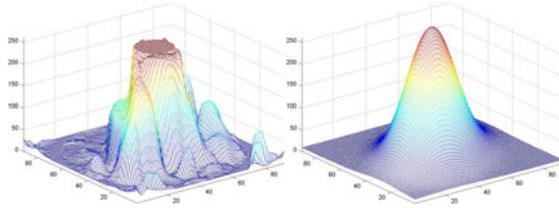


Fig. 4. Left: The grayscale projection data of the wheel in Figure 3 using KFD projection with RBF kernel and $\sigma = 1$. Right: The final result after nonlinear LS fit.

Figures 3 (f-g) illustrate the effect of the bandwidth parameter σ^2 . Figure 3 (f) uses the bandwidth selected using so called Silverman's rule of thumb [12], widely used in kernel density estimation and defined by $\hat{\sigma}_{\text{rot}} = 1.06\hat{\sigma}_x N^{-\frac{1}{5}}$, where $\hat{\sigma}_x$ is the sample standard deviation of the data and N is the data length. In the example in Figure 3 (f) the rule of thumb gives $\sigma_{\text{rot}} = 1.37$. Figure 3 (g) on the other hand illustrates the result with fixed $\sigma = 0.5$.

3 Elliptic Shape Model

The transition from the background (the healthy skin) to the foreground (the wheel) can be quite smooth, and the KFD-projected image may contain several individual foreground regions although the image has only one wheel. This is mostly due to the noise in the data, whose effect is greatly emphasized by the grayscale projection. Therefore, simple thresholding results in ragged boundaries, and unrealistic wheel size estimates. In order to increase the robustness of the segmentation, we fit a shape model for the appearance of the wheel. Since the manual measurement assumes that the wheals are ellipses, an elliptic shape model seems reasonable. Thus, the problem is to find an ellipse that divides the image into two maximally inhomogeneous areas.

Since there are an infinite amount of ellipses, we have to limit the search space somehow. This can be done by fitting a model to the grayscale projection and considering only the isosurfaces of the model. Based on Figure 3, the Gaussian surface seems an appropriate model for the spatial grayscale distribution in this case. Moreover, it suits our assumption of elliptic wheals, because the isosurfaces of the two-dimensional Gaussian are ellipses.

More specifically, the Gaussian model is defined by

$$f(\mathbf{x}; c, \mathbf{x}_0, \Sigma) = c \cdot \exp(-(\mathbf{x} - \mathbf{x}_0)^T \Sigma (\mathbf{x} - \mathbf{x}_0)), \quad (6)$$

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

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

$$\min_{c, \Sigma, \mathbf{x}_0} \sum_{k=0}^N (z_k - f(\mathbf{x}_k; c, \mathbf{x}_0, \Sigma))^2, \quad (7)$$

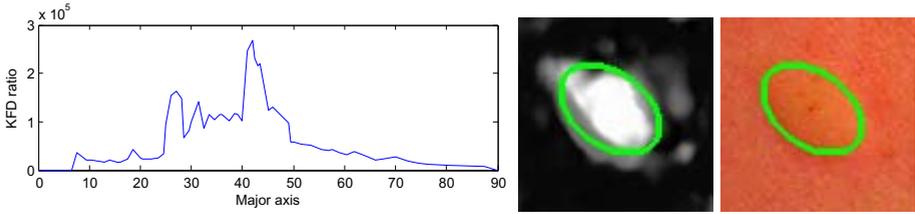


Fig. 5. Left: The KFDR as a function of the ellipse size. Center: The maximally separating ellipse overlaid on top of the corresponding KFD projection. Right: The maximally separating ellipse overlaid on top of original RGB data.

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

Fitting the Gaussian is a nontrivial problem, although lot of literature on the topic exists (e.g., [1]). However, the easiest approach is to use software packages such as Matlab Optimization toolbox to find the optimal parameters. In order to avoid local minima, we initialized the iterative search with the parameters obtained from a logarithmic transformation of the model and the data. This makes the problem linear least squares, and provides a good starting point. Figure 4 shows the original grayscale data on the left, the result of logarithmic fitting in the center and the result of nonlinear iterative fitting on the right.

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

The definition of a good ellipse among the candidates needs some measure of separation between the segmented areas. Recent work by Harchaoui *et al.* [6] considers using the Kernel Fisher Discriminant Ratio (KFDR) for testing the homogeneity between two sets, which coincides well with our use of KFD for grayscale projection in Section 2.

In other words, we test all ellipses that are cross sections of the fitted Gaussian and attempt to maximize the KFDR of Eq. (3) with respect to training sets defined by the ellipse. The situation is similar to the grayscale projection, but now we are not looking for a good classifier for the RGB data, but only assessing how well the data *could be classified*. Unlike Section 2, the choice of the training samples is now based on the boundaries of the ellipse to be tested. Note that this is not equivalent to calculating the variances directly from the projections of Figure 3, because the projection is calculated separately for the training sets determined by each ellipse candidate.

Sometimes the KFDR separability criterion results in very small ellipses, because a small foreground training set tends to be well separable. As an extreme example, an ellipse containing only a single pixel has extremely good separability assuming no other pixel has exactly the same RGB value. Thus, we decided to

modify the criterion by multiplying it with the cardinality of the smaller training set. Alternatively, we could set a minimum size restriction for the ellipse.

An example of the separability test is shown in Figure 5. The figure shows the KFDR between the "inside" class and the "outside" class for ellipses with different radius. It can be seen that the maximal separation is obtained at radius 42, and the corresponding ellipse is illustrated in Figure 5, as well.

4 Experimental Results

The results from the described method are compared to manual wheal segmentations (made by a non-medical expert). The similarity measure used by Celebi et al. [2] compares the *areas* of the segmentations. For our purposes, this is not an appropriate criterion, since ultimately we are interested in the major and minor axes of the wheal. The error in areas increases quadratically with respect to the axes, which is not desirable. Instead, we used the following error criterion between the computer segmentation A and the manual ground truth B :

$$E(A, B) = \frac{\sqrt{\text{Area}(\text{OR}(A, B))} - \sqrt{\text{Area}(\text{AND}(A, B))}}{\sqrt{\text{Area}(B)}}, \quad (8)$$

where $\text{OR}(A, B)$ consists of pixels segmented as foreground in A or B , and $\text{AND}(A, B)$ of foreground pixels in both A and B . Moreover, $\text{Area}(A)$ is the number of foreground pixels in A . The favourable property of Eq. (8) is that it increases linearly with respect to the error in major and minor axes. For example, it can be shown that the error measure for concentric circles with radii $r + a$ and $r - a$ are equal if the true radius is r . This is not the case with the error of [2].

Examples of segmentation results are illustrated in Figure 6. The figure shows the result of manual segmentation (red) compared with the result of the proposed

Table 1. The comparison of automated wheal measurement methods in terms of the error of Eq. (8). Each row defines an initial color space, each column corresponds to a grayscale transformation. The last column corresponds to a manually designed transformation based on what looks good. In the RGB case, the *Ad Hoc* transformation is the difference $G - B$, in the $L^*a^*b^*$ case it is the a^* component, and in the HSV case the H component.

		<i>Gaussian</i> ($\sigma = 0.5$)	<i>Gaussian</i> ($\sigma = 1$)	<i>Gaussian</i> ($\sigma = \sigma_{\text{rot}}$)	<i>Linear</i> <i>kernel</i>	<i>2. order</i> <i>kernel</i>	<i>Celebi</i> <i>method</i>	<i>Ad</i> <i>hoc</i>
With shape model	<i>RGB</i>	0.3283	0.2152	0.2331	0.4008	0.2119	1.5159	0.2552
	$L^*a^*b^*$	0.2120	0.2296	0.2141	0.1853	0.2154	1.4405	0.9945
	<i>HSV</i>	0.2501	0.2225	0.1951	0.4922	0.3948	1.2466	0.6302
Without shape model	<i>RGB</i>	0.4850	0.2327	0.1871	0.2602	0.2170	1.1963	0.2799
	$L^*a^*b^*$	0.2274	0.1866	0.3017	0.2403	0.2190	1.0266	1.4060
	<i>HSV</i>	0.2218	0.2219	0.5129	0.3932	0.3435	1.0099	0.4573

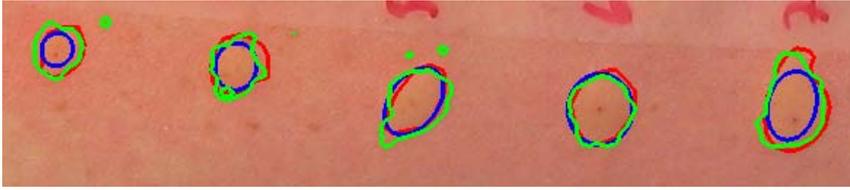


Fig. 6. An example of segmentation result. The red boundary is the result of manual segmentation, while blue and green boundaries represent the result of our method with and without the elliptic shape model, respectively. The errors (with / without shape model) for the five wheals are as follows (from left to right): 0.3144/0.2157, 0.2284/0.1562, 0.1130/0.2336, 0.1233/0.1269, 0.2457/0.1212.

method with (blue) and without (green) the shape model. Table 1 represents the average errors with different grayscale transformations. The test data consists of seven wheals including those shown in Figure 6. The five first columns represent different KFD projections designed using training data, while in the last two columns the projection is designed in an unsupervised or *ad hoc* manner.

From the results one can clearly see that the KFD projections have the smallest errors. The errors are many times larger in the last two columns. On the other hand, the best performing projections (e.g., the Gaussian kernel) seems to make the elliptic shape model unnecessary. On the other hand, this is rather obvious when looking at, e.g., the result in Figure 3(f), where the separation between foreground and background is very clear.

Another reason for a worse than expected performance of the shape model is seen in Figure 6. The manually segmented wheals are not ellipses, so the elliptic model can not reach zero error even in theory. The best cases are the ones where the true wheal is ellipse-shaped with no elongated pseudopodia, i.e., 3rd and 4th wheals from the left. In all other cases the wheal shape is more irregular, and the shape model results in the largest inscribed ellipse. However, there is some randomness in the results due to the small N . We plan to study the performance with larger N and compare them with the manual results of a trained physician.

5 Conclusions

In this paper we proposed a method for automatic segmentation of allergic reactions. The method combines an optimal grayscale transformation with an elliptic shape model for the allergic reactions. Experiments show that the method can efficiently quantify the size of the wheal. The experiments also indicate that the grayscale transformation is sometimes powerful enough to render the elliptic shape mode unnecessary. However, the shape model results may be in coherence with medical doctor's measurements, since they also disregard the pseudopodia.

We plan to continue the development of the algorithm into several directions. One is to consider a recently introduced method of *graph cuts* with shape priors. It will also be interesting to compare the results with those of a medical

doctor, together with larger population of test persons. Also the robustness of the method to, e.g., the starting point pointed by the user is an important topic for practical applicability. Additional topics of future work include automatic wheal detection and extension of the method to time series measured from video.

References

1. Brändle, N., Chen, H., Bischof, H., Lapp, H.: Robust parametric and semi-parametric spot fitting for spot array images. In: 8th Intl. Conf. on Intell. Syst. for Mol. Biol., ISMB 2000, pp. 1–12 (2000)
2. Celebi, M., Iyatomi, H., Schaefer, G.: Contrast enhancement in dermoscopy images by maximizing a histogram bimodality measure. In: 16th IEEE Int. Conf. on Image Proc. (ICIP), pp. 2601–2604 (2009)
3. Fisher, R.: The use of multiple measurements in taxonomic problems. *Annals of Eugenics* 7, 179–188 (1936)
4. Gomez, D.D., Clemmensen, L.H., Ersbøll, B.K., Carstensen, J.M.: Precise acquisition and unsupervised segmentation of multi-spectral images. *Comp. Vis. and Image Understanding* 106(2-3), 183–193 (2007)
5. Gomez, D., Butakoff, C., Ersboll, B., Stoecker, W.: Independent histogram pursuit for segmentation of skin lesions. *IEEE Trans. Biomed. Eng.* 55(1), 157–161 (2008)
6. Harchaoui, Z., Bach, F., Eric, M.: Testing for homogeneity with kernel fisher discriminant analysis. In: *Adv. in Neural Inf. Proc. Syst.* 20, pp. 609–616. MIT Press, Cambridge (2008)
7. Mika, S., Ratsch, G., Weston, J., Scholkopf, B., Mullers, K.: Fisher discriminant analysis with kernels. In: *Proc. IEEE Neural Netw. for Signal Process.* IX, pp. 41–48 (1999)
8. Nischik, M., Forster, C.: Analysis of skin erythema using true-color images. *IEEE Trans. Med. Imag.* 16(6), 711–716 (1997)
9. Roullot, E., Autegarden, J.-E., Devriendt, P., Leynadier, F.: Segmentation of erythema from skin photographs for assisted diagnosis in allergology. In: Singh, S., Singh, M., Apte, C., Perner, P. (eds.) *ICAPR 2005. LNCS*, vol. 3687, pp. 754–763. Springer, Heidelberg (2005)
10. Santos, R., Mlynek, A., Lima, H., Martus, P., Maurer, M.: Beyond flat weals: validation of a three-dimensional imaging technology that will improve skin allergy research. *Clin. Exp. Dermatol.* 33(6), 772–775 (2008)
11. Schölkopf, B., Smola, A.J.: *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, 1st edn. MIT Press, Cambridge (2001)
12. Silverman, B.: *Density Estimation for Statistics and Data Analysis*. Chapman-Hall, Boca Raton (1986)
13. Wöhrle, S., Vigl, K., Binder, M., Stingl, G., Prinz, M.: Automated measurement of skin prick tests: an advance towards exact calculation of wheal size. *Experimental Dermatology* 15(2), 119–124 (2006)