# Quasiconvex Alignment of Multimodal Skin Images for Quantitative Dermatology

Siddharth K. Madan
Rutgers University
94 Brett Road,
Piscataway,
NJ, 08854.
smadan@eden.rutgers.edu

Kristin J. Dana
Rutgers University
94 Brett Road,
Piscataway,
NJ, 08854.
kdana@ece.rutgers.edu

Oana G. Cula
Johnson & Johnson
199 Grandview Road,
Skillman,
NJ, 08558.
gcula@cpcus.jnj.com

## Abstract

*In quantitative dermatology, high resolution sensors provide images that capture fine scale features like pores, birthmarks, and moles. Breathing and minute movements result in misregistration of micro level features. Many computer vision methods for dermatology such as change detection, appearance capture, and multi sensor fusion require high accuracy point-wise registration of micro level features. However, most computer vision algorithms are based on macro level features such as eyes, nose, and lips, and aren't suitable for registering micro level features. In this paper, we develop a practical robust algorithm to align face regions using skin texture with mostly indistinct micro level features. In computer vision, these regions would typically be considered featureless regions. Our method approximates the face surface as a collection of quasi-planar skin patches and uses quasiconvex optimization and the $L_\infty$ norm for estimation of spatially varying homographies.*

*We have assembled a unique dataset of high resolution dermatology images comprised of over* 100 *human subjects. The image pairs vary in imaging modality (crossed, parallel and no polarization) and are misregistered due to the natural non-rigid human movement between image capture. This method of polarization based image capture is commonly used in dermatology to image surface and subsurface structure. Using this dataset, we show high quality alignment of "featureless" regions and demonstrate that the algorithm works robustly over a large set of subjects with different skin texture appearance, not just a few test images.*

## 1. Introduction

With the growing popularity of high resolution sensors, skin images with fine scale texture details such as pores, moles, and birth marks are becoming increasingly common in quantitative dermatology. Algorithms focused on fine scale features are being developed. For example, in [20] the authors develop an algorithm to localize fine scale features like moles and birthmarks. Fine scale skin texture was also used in [5] to recognize skin type using a texton-model from images acquired under multiple illumination and viewing directions. High resolution face images always have some amount of misregistration due to breathing and other involuntary movement. While many excellent algorithms exist for face alignment, including face morphing [2], active appearance models [4, 18], and 3D deformable models [9], most registration algorithms deal with registration of distinct macro level features like eyes, nose, and lips, as opposed to fine scale registration. Registration algorithms based on macro level features may be ideal for video face tracking, but dermatology applications using very high resolution images require alignment of face regions so that even skin texture details such as pores, spots and lines are aligned.

In this paper we develop a registration algorithm for registering micro level features in high resolution face images. Face images are acquired under three different modalities: cross polarized, parallel polarized, and without using any polarizers. The different modalities enhance different types of skin features. Because of involuntary, human movement misregistration exists in the multimodal images acquired sequentially in time. We approximate the face as a collection of quasi-planar patches, register the patches, and use a piecewise planar assumption to register large regions of skin data. Registration is accomplished using feature matching and quasiconvex optimization. The quasi-planar patches do not contain any macro level features, so running interest point detector directly on the quasi-planar patches detects only a few feature points. In order to detect a large number of feature points, we apply an intensity transformation to the quasi-planar patches and detect the micro level features in the enhanced patches. Unlike optimization algorithms often encountered in computer vision algorithms, the optimization algorithm solved in our case has only one minimum, and the minimum can be efficiently found. We have tested our algorithms on a large dataset comprising of over 100

978-1-4244-3993-5/09/$25.00 ©2009 IEEE

human subjects. High quality registration results obtained for large number of images indicates our algorithm works well on general images, and not just on few test images. [1]

The remaining paper is organized as follows. In section 2 we discuss the interaction of skin surface with visible light, and the importance of multimodal imaging in dermatology. In section 3, we discuss the acquisition of high resolution multimodal skin image database of 100 subjects. In sections 4–6 we discuss the feature extraction from featureless quasi-planar regions, detection and removal of gross outliers, and registration of quasi-planar patches. We discuss the registration of larger skin regions in section 7, and present experimental results in section 8.

## 2. The surface-subsurface reflectance model

Skin is a multi-layered medium, with complex fine scale geometry, an oily layer at the air-skin interface, and with layers of different types of cells in the stratum corneum, epidermis, and dermis. Consequently, the way skin reflects visible light is defined by the reflection and inter-reflection of incident light at the interfaces between layers with different indices of refraction. A simplified model of skin reflectance is the surface-subsurface model [19, 6], which models the reflectance as a sum of surface and subsurface components. The surface component is the part of the incident light reflected off the skin surface. The surface component arises due to the oily layer present on normal facial skin. The subsurface component is the part of the incident light traveling through the stratum corneum and the epidermis. The subsurface component suffers multiple subsurface scattering events before it exits the skin. Figure 1 schematically illustrates the multi-layered structure of skin and its interaction with visible light.

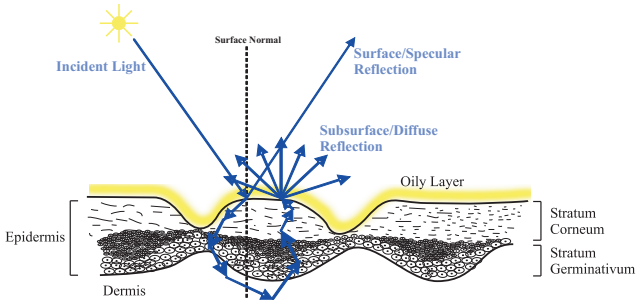



Figure 1. Interaction of light with the skin surface.

When the skin is illuminated with linearly polarized light, the surface reflection preserves the polarization of the incident light, but the polarization status of the subsurface component gets randomized by the birefringent dermal collagen fibers [1]. The dichotomy in polarization of the reflected components makes linearly polarized imaging of skin quite useful in dermatology [22, 12]. The incident light can be polarized by placing a linear polarizer in front of the light source. If a polarizer is placed in front of the sensor, the intensity measured by the sensor depends on the relative orientation of the polarizer in front of the light source and the polarizer in front of the sensor [13]. If the polarizer in front of the sensor is parallel to the polarizer in front of the light source, the entire surface component is measured by the sensor, but only half the subsurface component reaches the sensor. If $I_s$ denotes the surface component and $I_d$ denotes the subsurface component, the intensity $I_p$ measured by the sensor is,

$$I_p = I_s + \frac{1}{2} I_d. \tag{1}$$

We use the term *parallel polarized image* for an image acquired with parallel polarizers in front of the light source and the sensor. A parallel polarized image enhances the surface component, and brings out surface features like raised borders of lesions, pore structure, and wrinkles [15]. If the polarizer in front of the sensor is perpendicular to the polarizer in front of the light source, the entire surface component gets blocked, and the sensor measures only half the subsurface component. The intensity $I_x$ measured by the sensor is,

$$I_x = \frac{1}{2} I_d. \tag{2}$$

We use the term *cross polarized image* for an image acquired with perpendicular polarizers in front of the light source and the camera. The cross polarized image brings out subsurface skin features like color variation due to melanin erythema [16]. The subsurface component penetrates no more than about 300 micrometers below the skin surface, therefore the subsurface component offers a reasonable measurement of melanin and superficial blood vessels residing in the epidermis and the upper part of the dermis, respectively. When no polarizers are placed in front of the light source and the camera, the intensity $I_v$ measured by the sensor would be the sum of the surface and the subsurface components.

$$I_v = I_s + I_d. \tag{3}$$

We use the term *visual image* for an image acquired without any polarizers in front of the sensor and the light source.

## 3. Multimodal skin image database

For the experimental results of this paper, we have created a new database for dermatology images. The image acquisition phase was done in a skin biology research center using custom equipment for skin imaging. We have imaged 100 subjects of different ethnicities: Caucasians, African Americans, South Asians, and East Asians. Each subject was imaged under three different modalities and from the left, right, and frontal view points. The illumination direction remains fixed with respect to the subject position. Figure 2 illustrates the imaging system used for data acquisition. The imaging system consists of a light source, a

[1] In the supplementary material, we have included a large number of registration results. The misregistration between the input skin regions, and the quality of registration can be best seen by viewing the images on a computer screen, one after the other, in quick succession.

sensor, and polarizers placed in front of the sensor and the light source. The sensor is a Canon Digital Rebel XTI camera equipped with a Canon 35 $f2.0$ lens and a Broncolor Picolite flash light. The sensor and the orientation of the polarizers are controlled automatically by a computer, which enables quick acquisition of the cross polarized, parallel polarized, and visual images. Every image is $2602 \times 3906$ in size. Figure 3 shows example images of four subjects in our database. The eye region has been blacked out to preserve the identity of the subjects. Even though the multimodal images were acquired rapidly and a chinrest was used, the acquired images are slightly misregistered. The misregistration arises due to small involuntary movements like breathing which are always present. Misregistration between multimodal images is a problem in quantitative dermatology applications for which an exact point to point registration is required. For example, in [23] the authors use polarized light imaging to delineate the border of skin cancer tumors, which requires an exact point to point registration between the acquired multimodal images. In [21], the authors use multimodal images to estimate the concentrations of oxy-hemoglobin, deoxy-hemoglobin, and water at each pixel location, which requires an exact point to point registration between the multimodal images. The misregistration between the images in the multimodal skin image database is nonrigid, so simple global motion models are not adequate. In this paper we propose a novel algorithm to register skin patches extracted from the multimodal face images. We register larger regions of the face by assuming a piecewise planar face model so that patch images can be stitched together.

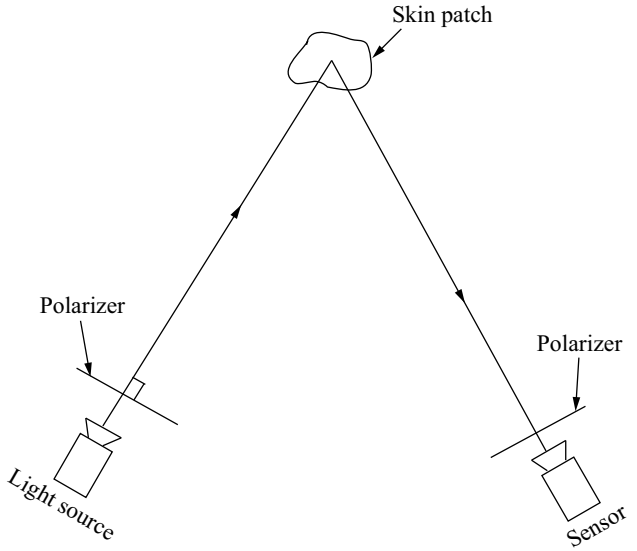



Figure 2. Imaging system to acquire polarized images.

Figure 3. Images of four subjects under three different imaging modalities. For each subject the left image in the set is the cross polarized image, the parallel polarized image is in the center, and the visual image is on the right.

## 4. Feature extraction and matching in featureless regions

Figure 4 (a), (b), and (c) show $2602 \times 3906$ sized cross polarized, visual, and parallel polarized images of a human face. The eye region has been blacked out to preserve the identity of the person being imaged. Figure 4 (d), (e), and (f) show $400 \times 400$ size patches extracted, around a common pixel location, from each imaging modality. The patches extracted are very small compared to the size of the face images, and can be approximated as quasi-planar surfaces. Note that the patches are extremely smooth and do not appear to contain high number of feature points. In our experiments the visual patch is the reference image, and both the cross polarized and the parallel polarized patches are registered onto the visual patch.

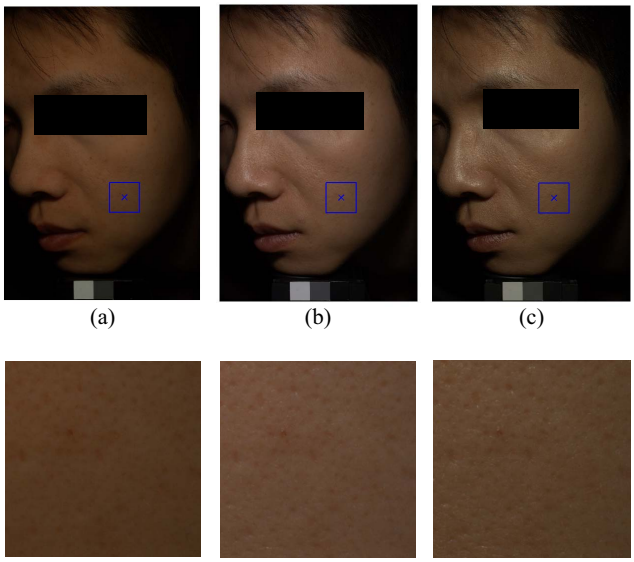
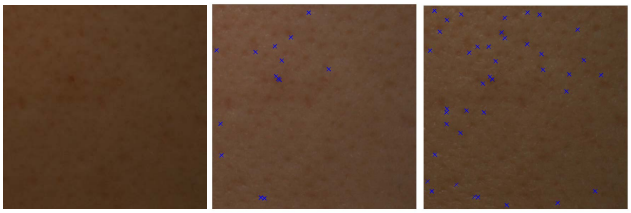



Figure 4. Patches extracted from the face images. (a), (b), and (c) are the cross polarized, visual, and the parallel polarized face images. (d), (e), and (f) are the patches extracted from the cross polarized, visual, and the parallel polarized images respectively. In the face images, '$\times$' denotes the common pixel location around which the patches have been extracted, and the rectangle denotes the boundary of the patch.

In order to register the cross polarized and parallel polarized patches onto the visual patch, we detect and match feature points in the patch images using the scale invariant feature transform (SIFT) [17]. Figure 5 shows the feature points detected when the SIFT detector is run on the cross polarized, visual, and parallel polarized patches. Note that running SIFT on the unprocessed patch images results in very few feature points being detected. The SIFT algorithm



(a) (b) (c)

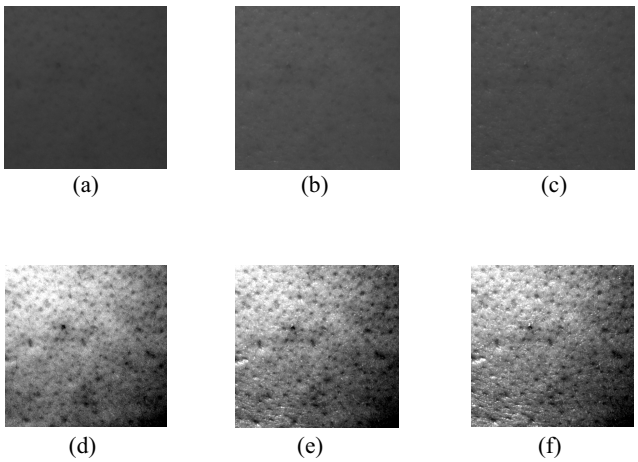
Figure 5. Feature points detected when SIFT is run on the patch images. (a): Feature points detected in the cross polarized patch. No feature points were detected in the cross polarized patch image. (b): Feature points detected in the visual patch. (c): Feature points detected in the parallel polarized patch.

was not able to detect any feature point in the cross polarized patch image. It could detect only $14$ points in the visual patch, and $40$ points in the parallel polarized patch image. Instead of using fiducial markers to assist registration, we apply a local intensity transform to enhance subtle features that are present in high resolution face images. We increase the contrast of the patch images by linearly stretching the intensities of the images, in grayscale, such that one percent of the intensities are saturated. Figure 6 shows the original and the contrast enhanced patch images in grayscale. Features are clearly brought out in the contrast enhanced patch images. To detect a large number of feature points, we run SIFT on the contrast enhanced grayscale images. Figure 7 shows the feature points obtained when SIFT was run on the contrast enhanced patch images. A total of $1465$ feature points were detected in the cross polarized patch, $2215$ feature points were detected in the visual patch, and $2452$ feature points were detected in the parallel polarized patch. The point set shown in Figure 7 are far too dense. However, a meaningful reduction of features can be accomplished.

SIFT describes every detected feature point by a vector, called the descriptor vector of the feature point, giving the gradient orientations in a region around the detected feature point. In the matching stage, SIFT removes the weak feature points, and matches a feature point $A$ to feature point $B$ if the distance between descriptors of A and B is less than a threshold $T$ given by,

$$T = k\Delta, \tag{4}$$

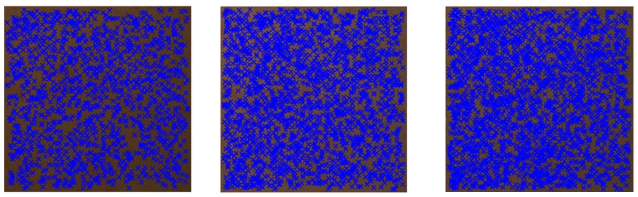
where $\Delta$ is the distance between the descriptor of $A$ and the descriptor of any other feature point, and $k$ is a threshold factor such that $k \in [0, 1]$. The number of matches obtained is sensitive to the value of the threshold factor $k$. Increasing the value of $k$ increases the number of matches obtained, but



(a) (b) (c)

(d) (e) (f)

Figure 6. The original and contrast enhanced patch images. (a), (b), and (c) show the original cross polarized, visual, and parallel polarized patch images in grayscale. (d), (e), and (f) show the grayscale contrast enhanced cross polarized, visual, and parallel polarized patch images. Note that the smooth skin region has abundant features for alignment in the contrast enhanced patch images.



(a) (b) (c)

Figure 7. Feature points detected when SIFT was run on the contrast enhanced patch images. (a): Feature points detected in the cross polarized patch. (b): Feature points detected in the visual patch. (c): Feature points detected in the parallel polarized patch.

a lot of the detected matches would be outliers. In our registration algorithm, we set $k = 1$ to obtain a large number of matches and remove the outliers in two stages as described in section 5.

## 5. Two stage detection and removal of outliers

In this section we discuss the removal of incorrect matches. We remove the incorrect matches in two stages. The two stage approach greatly increases the efficiency of the outlier removal step.

In the first stage, we remove matches which are easily detected as outliers. Misregistration in the images of the multimodal skin image database is due to breathing and other involuntary motion. This misregistration is expected to be small, but significant. Consequently, if the difference in the pixel locations of matched pair of feature points is greater than a certain threshold, we can conclude that the matched pair is an outlier. In particular, if the difference in the pixel location in the $x$ and $y$ directions are greater than a certain threshold $\Delta x$ and $\Delta y$ respectively, we conclude that

the matched pair is an outlier pair. We refer the matches for which the difference in pixel locations is greater than the $\Delta x$ and $\Delta y$ threshold as *gross outliers*. A pair of matched feature points with pixel locations $(x_1, y_1)$, and $(x_2, y_2)$, in the two patch images to be registered, are gross outliers if,

$$|x_1 - x_2| \geq \Delta x, \tag{5}$$
$$|y_1 - y_2| \geq \Delta y. \tag{6}$$

We select $\Delta x$ and $\Delta y$ empirically.

In the second stage, the remaining outliers are removed using RANSAC [7, 8]. Using RANSAC, an initial estimate of homography [11] between the cross polarized and visual patches, and between the parallel polarized and visual patches is obtained. Removing the easily detected outliers before running RANSAC increases the probability of a randomly chosen match being an inlier match, which reduces the number of iterations required by RANSAC. Reduction in the number of required iterations increases the efficiency of the outlier removal step.

Figure 8 shows the feature points retained after removing all the outliers. Figure 8 (a) and (b) show the feature points retained in the cross polarized and visual patch images, when the cross polarized patch is registered onto the visual patch image. A total of 96 feature points were retained. In Figure 8 (e) the feature points are plotted together. Green points indicate the pixel locations in the cross polarized patch image and the red points indicate the pixel locations in the visual patch image. The difference in the locations of the green and red points indicates the cross polarized and visual patch images are misregistered nonrigidly. Note that the extent of misregistration is small. However, the misregistration is significant for quantitative dermatology applications where precise point to point registration is required. Figure 8 (c) and (d) show the feature points retained in the parallel polarized and visual patch images, when the parallel polarized patch is registered onto the visual patch image. A total of 491 feature points were retained. In Figure 8 (f) the feature points are plotted together. To maintain clarity only 123 of the 491 feature points are plotted in Figure 8 (f). Green points indicate the pixel locations in the parallel polarized patch image and the red points indicate the pixel locations in the visual patch image. The difference in the locations of the green and red points indicates the parallel polarized and visual patch images are misregistered.

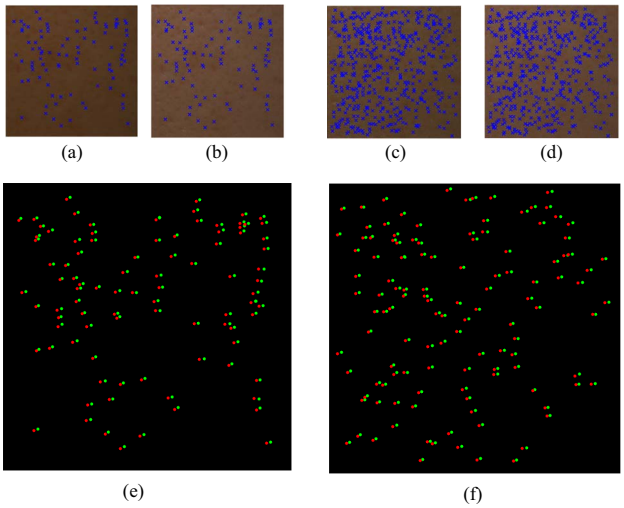



Figure 8. (a) and (b) show the feature points retained during registration of the cross polarized patch onto the visual patch image. (c) and (d) show the feature points retained during the registration of the parallel polarized patch onto the visual patch. In (e), the feature points retained during the registration of the cross polarized and visual patches, are plotted together. The pixel locations in the cross polarized patch are denoted in green and the the pixel locations in the visual patch are denoted in red. The difference in the locations indicate the patch images are misregistered. In (f), 123 of the 491 feature points retained during the registration of the parallel polarized and visual patches, are plotted together. The pixel locations in the parallel polarized patch are denoted in green and the pixel locations in the visual patch are denoted in red.

## 6. Quasiconvex optimization based registration of quasi-planar patches

The patch images extracted from the face are quasi-planar surfaces; hence a homography can be estimated between them. We register the cross polarized patch onto the visual patch by estimating a homography between the cross polarized and visual patch images, and we register the parallel polarized patch onto the visual patch image by estimating a homography between the parallel polarized and visual patch images. We estimate the homographies using the quasiconvex optimization approach recently introduced in [14]. The quasiconvex approach minimizes the $L_\infty$ norm of the reprojection errors. The advantage of the quasiconvex approach is that the cost function minimized has only one minimum [10], and the minimum can be efficiently found using the bisection algorithm [3]. Previously the $L_\infty$ approach was presented with a small number of experimental results. Here we demonstrate its utility on a larger dataset and on non rigid motions approximated with piecewise homographies.

Using the point correspondences we estimate a matrix $\mathbf{A}_c$, which relates the homogeneous coordinate of a point in the visual image to the homogeneous coordinate of the point in the misregistered cross polarized image as follows,

$$s \begin{bmatrix} \mathbf{x}_c \\ 1 \end{bmatrix} = \mathbf{A}_c \begin{bmatrix} \mathbf{x}_v \\ 1 \end{bmatrix}, \tag{7}$$

where $s$ is a scale factor, $\mathbf{x}_c = [p_c \;\; q_c]^T$ and $\mathbf{x}_v = [p_v \;\; q_v]^T$ are the pixel locations of the point in the misregistered cross polarized and visual patch images. Similarly we estimate a matrix $\mathbf{A}_p$ which relates the homogeneous coordinate of a point in the visual image to the homogeneous coordinate of the point in the misregistered parallel polarized image. In

this discussion, we give the estimation of $\mathbf{A}_c$. The estimation process of $\mathbf{A}_p$ is identical. The discussion in current section is based on [14].

Let $\mathbf{x}_c^{(i)} = [p_c^{(i)} \quad q_c^{(i)}]^T$ and $\mathbf{x}_v^{(i)} = [p_v^{(i)} \quad q_v^{(i)}]^T$ denote the pixel locations of the $i^{th}$ feature point in the misregistered cross polarized and visual patch images, and let $\mathbf{A}_c = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3]^T$. The projection, $\hat{\mathbf{x}}_c^{(i)}$, of $\mathbf{x}_v^{(i)}$ onto the misregistered cross polarized patch is given by,

$$s \begin{bmatrix} \hat{\mathbf{x}}_c^{(i)} \\ 1 \end{bmatrix} = \mathbf{A}_c \begin{bmatrix} \mathbf{x}_v^{(i)} \\ 1 \end{bmatrix}. \tag{8}$$

From 8, the coordinate $\hat{\mathbf{x}}_v^{(i)}$ is given by,

$$\hat{\mathbf{x}}_c^{(i)} = \begin{bmatrix} \mathbf{a}^{(i)^T}\mathbf{r}/\mathbf{c}^{(i)^T}\mathbf{r} \\ \mathbf{b}^{(i)^T}\mathbf{r}/\mathbf{c}^{(i)^T}\mathbf{r} \end{bmatrix}, \tag{9}$$

where,

$$\mathbf{a}^{(i)} = \begin{bmatrix} \mathbf{x}_v^{(i)^T} & 1 & \mathbf{0}_{1\times 3} & \mathbf{0}_{1\times 3} \end{bmatrix}^T,$$
$$\mathbf{b}^{(i)} = \begin{bmatrix} \mathbf{0}_{1\times 3} & \mathbf{x}_v^{(i)^T} & 1 & \mathbf{0}_{1\times 3} \end{bmatrix}^T,$$
$$\mathbf{c}^{(i)} = \begin{bmatrix} \mathbf{0}_{1\times 3} & \mathbf{0}_{1\times 3} & \mathbf{x}_v^{(i)^T} & 1 \end{bmatrix}^T,$$
$$\mathbf{r} = \begin{bmatrix} \mathbf{r}_1^T & \mathbf{r}_2^T & \mathbf{r}_3^T \end{bmatrix}^T.$$

The error in reprojection is,

$$\begin{aligned} f_i(\mathbf{r}) &= \|\mathbf{x}_c^{(i)} - \hat{\mathbf{x}}_c^{(i)}\|_2 \\ &= \frac{\|\mathbf{e}^{(i)}\|_2}{\mathbf{c}^{(i)^T}\mathbf{r}}, \end{aligned} \tag{10}$$

where the vector $\mathbf{e} = [e_1 \ \ e_2]^T$ is,

$$e_1 = (p_c^{(i)}\mathbf{c}^{(i)} - \mathbf{a}^{(i)})^T\mathbf{r},$$
$$e_2 = (q_c^{(i)}\mathbf{c}^{(i)} - \mathbf{b}^{(i)})^T\mathbf{r}.$$

The optimal $\mathbf{r}$ is defined as the solution of the optimization problem,

$$\begin{aligned} &\text{minimize} \quad F(\mathbf{r}) = \max_i f_i(\mathbf{r}) \\ &\text{subject to} \quad \mathbf{c}^{(i)^T}\mathbf{r} > 0 \quad i = 1, \ldots, n. \end{aligned} \tag{11}$$

The optimal $\mathbf{r}$ minimizes the maximum reprojection error. The optimization problem defined in 11 is a quasiconvex optimization problem [14]. The objective function, $F(\mathbf{r})$, is a quasiconvex function, i.e., the domain of the objective function, and the sets $S_\alpha = \{\mathbf{r} \ : \ F(\mathbf{r}) \leq \alpha\}$ are convex sets. The global minimum of the optimization problem can be easily found using the bisection algorithm [3].

After estimating $\mathbf{A}_c$ and $\mathbf{A}_p$, the patches are registered using image warping. Consider a point $P$ with pixel location $[x \ y]^T$ in the registered cross polarized patch. To obtain the intensity at $P$, we project it onto the misregistered cross polarized patch using $\mathbf{A}_c$. Let $[x' \ y']^T$ be pixel location of $P$ when projected to the misregistered cross polarized patch. The intensity at point $P$ in the registered cross polarized patch is the intensity at $[x' \ y']^T$ in the misregistered cross polarized patch; the intensity at $[x' \ y']^T$ is obtained using bilinear interpolation. Similarly, we can use $\mathbf{A}_p$ to register the parallel polarized patch onto the visual patch.

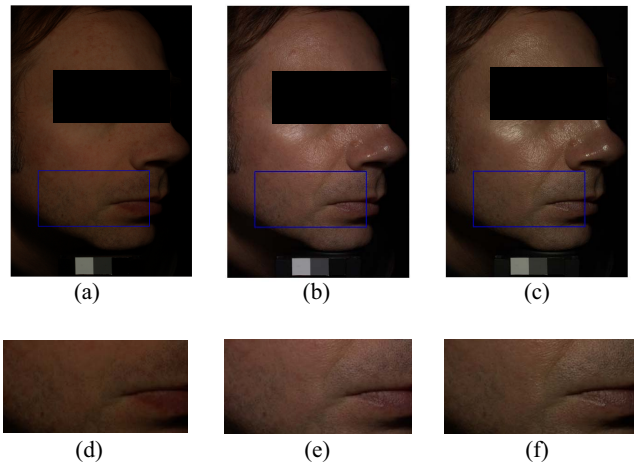


Figure 9. Face images and the extracted skin regions. (a), (b), and (c) show the cross polarized, visual, and parallel polarized face images. We register the 1600 sized skin region defined by the rectangle shown on the face images. We have blacked out the eye region to preserve the identity of the person. The extracted skin regions are shown in (d), (e), and (f).

## 7. Registering large regions of skin data through qausi-planar approximation

In the prior sections, we have discussed the registration of small quasi-planar skin patches. In this section, we discuss the registration of larger regions of skin data.

Figure 9 (a), (b), and (c) show $2602 \times 3906$ sized cross polarized, visual and parallel polarized images of a face. We register the $1600 \times 800$ sized skin region within the rectangle marked on the face images. Figure 9 (d), (e), and (f) show the skin data extracted from each imaging modality. The skin regions are eight times larger than the $400 \times 400$ sized quasi-planar patches shown in Figure 4, and are too big to be approximated as quasi-planar surfaces. We treat the visual skin region as the reference, and register the cross polarized and parallel polarized regions onto the visual skin region.

We divide the skin region extracted from each modality into smaller $400 \times 400$ sized component patches, where each patch is non-overlapping and regularly spaced. Each component patch can be approximated as a quasi-planar surface. We use feature extraction and quasiconvex optimization to register each component cross polarized or parallel polarized patch onto the corresponding component visual patch. Registration of the component patches results in a high resolution registration of the large polarized and visual skin

regions.

## 8. Experimental results

We have tested our algorithm on a dataset consisting of high resolution face images of over 100 subjects. We have included our complete set of registration results in the supplementary material. Both the misregistration in the input images, and the final registration results are best seen by viewing the images on a screen, one after the other, in quick succession. In the supplementary material, we have included videos showing the misregistered images one after the other, and the registered images one after the other. Excellent results obtained for a large number of images shows that our registration technique works well on general images, and not just on few test images. In this section, we present example registration results for the cross polarized and visual patches show in Figure 4, and the large parallel polarized and visual skin regions shown in Figure 9.

Figure 10 (a), (b), and (c) show the misregistered cross polarized patch image, the visual patch image, and the cross polarized patch registered onto the visual patch. In Figure 10 (d) the pixel locations of the corresponding features in the misregistered cross polarized and the visual patch images are plotted together. The green dots indicate the pixel location in the cross polarized image, and the red dots indicate the pixel location in the visual patch image. The pixel locations in the registered cross polarized and visual patch images are plotted together in Figure 10 (e). Note that the pixel locations concur, indicating that the patch images are well registered.

Figure 11 (a), (b), and (c) show the misregistered parallel polarized, visual, and the registered parallel polarized skin regions extracted from the face. In Figure 11 (d) pixel locations of 72 corresponding feature points in the misregistered parallel polarized and visual skin regions are plotted together. A total of 7130 feature points were detected, but inorder to maintain clarity only 72 feature points were plotted. The pixel locations in the parallel polarized region are plotted in green and the pixel locations in the visual skin region are plotted in red. In Figure 11 (e) the pixel locations of the 72 feature points in the registered parallel polarized and visual skin regions are plotted together. Note that the pixel locations concur, indicating that the skin regions are well registered.

## 9. Conclusion

In this paper, we have presented a novel algorithm to register high resolution skin data using micro level features and quasiconvex optimization. We emphasize that our approach is not simply a SIFT registration, but instead makes several novel contributions. Our approach of registering faces using *featureless* regions by enhancing micro level features is a new paradigm. Common face alignment algorithms in computer vision use large-scale features like the eyes, nose and mouth, but this approach is not sufficient for dermatology studies where precise alignment of pore-level features is needed. We've collected a large database of images for dermatology with a wide range of age and ethnic variations. Using these images, we provide an in-depth testing of a recent quasiconvex homography estimation approach using the $L_\infty$ norm. We also use this large image set to test the assumption that skin regions with non-rigid motion can be aligned using piecewise planar approximation.

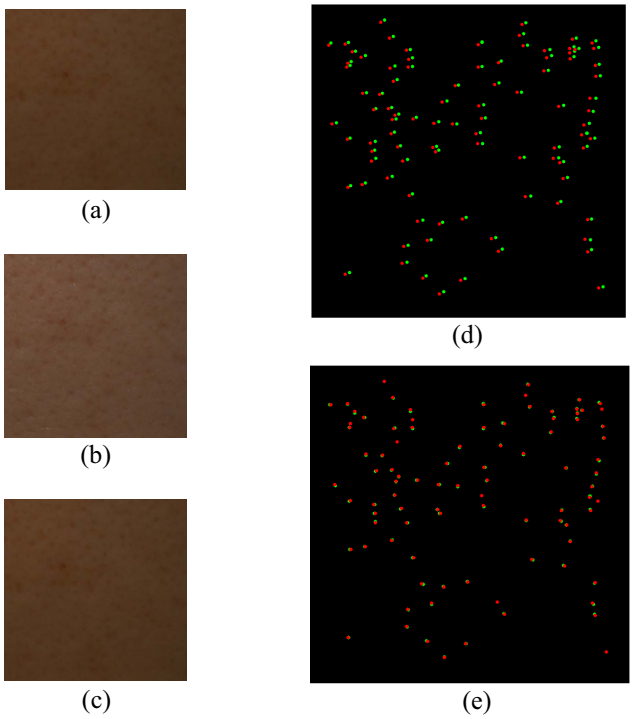



Figure 10. Registration of cross polarized and visual patch images. (a), (b), and (c) show the input cross polarized, visual, and the cross polarized registered to the visual patch images. In (d), the pixel locations of feature points in the misregistered cross polarized image and visual image are plotted together. In (e), the pixel locations of feature points in the registered cross polarized and visual patch images are plotted together.

Our registration algorithm is an important enabling step for quantitative dermatology using photography. The typical digital image is a recorded visual observation enabling only qualitative assessment. However, using alignment and illumination-controlled image capture, a digital camera can be employed as a scientific measurement device. Computer algorithms can process aligned skin images over time as a spatio-temporal signal in order to perform change detection, multi-sensor fusion or enhanced visual comparisons. As evident by the multimodal skin image database, non-rigid misregistration occurs even in images acquired in quick succession and with the use of chinrests. When image sets have larger levels of misregistration there are numerous algorithms in the computer vision literature which can first perform macro-scale face alignment using larger features such as eyes, nose and mouth. These algorithms can serve as

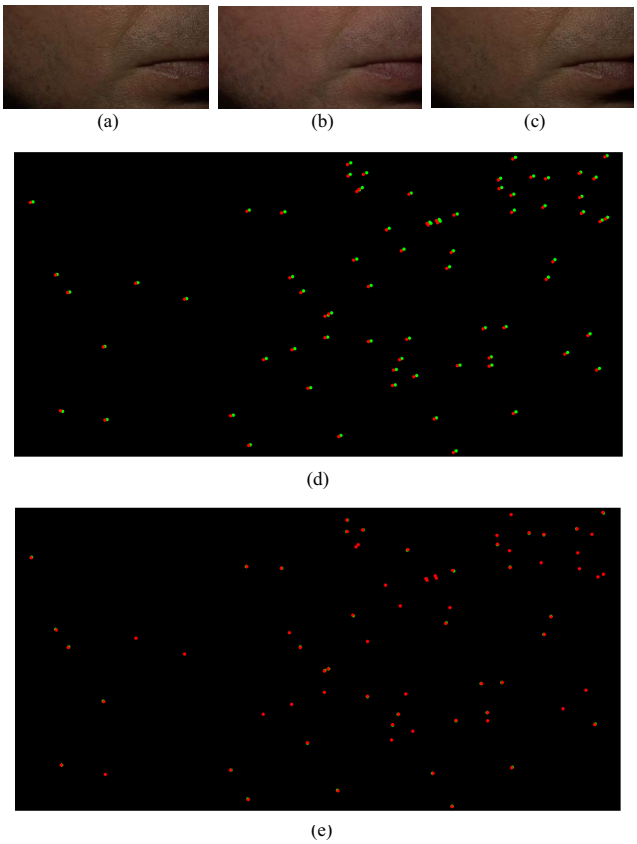


Figure 11. Registered parallel polarized and visual skin regions. (a), (b), and (c) show the misregistered parallel polarized, the visual, and the registered parallel polarized skin regions. In (d), pixel locations of 72 corresponding feature points in the misregistered parallel polarized and the visual skin regions are plotted together. A total of 7130 feature points were detected, but inorder to maintain clarity only 72 feature points were plotted. In (e) the pixel locations of the 72 corresponding feature points in the registered parallel polarized and visual skin regions are plotted together. The pixel locations in the parallel polarized skin region are indicated in green and the pixel locations in the visual skin region are indicated in red.

a pre-processor to the high-resolution alignment presented here that achieves point-to-point registration between the acquired images.

## References


[1] R. Anderson. Polarized light examination and photography of the skin. *Archives of Dermatology*, 127(7), 1991. 2

[2] V. Blanz and T. Vetter. A morphable model for the synthesis of 3D faces. In *ACM SIGGRAPH*, pages 187–194, 1999. 1

[3] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. 5, 6

[4] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 23(6):681–685, 2001. 1

[5] O. G. Cula, K. J. Dana, F. P. Murphy, and B. K. Rao. Skin texture modeling. *International Journal on Computer Vision*, 62(1–2):97–119, 2005. 1

[6] P. Debevec, T. Hawkins, C. Tchou, H. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *ACM SIGGRAPH*, pages 145–156, 2000. 2

[7] 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. 5

[8] D. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, 2002. 5

[9] L. Gu and T. Kanade. 3D alignment of face in a single image. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 1305–1312, 2006. 1

[10] R. Hartley and F. Schaffalitzky. L$\infty$ minimization in geometric reconstruction problem. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 504–509, 2004. 5

[11] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000. 5

[12] S. L. Jacques, J. C. Ramella-Roman, and K. Lee. Imaging superficial tissues with polarized light. *Lasers in Surgery and Medicine*, 26(2):119–129. 2

[13] S. L. Jacques, J. C. Ramella-Roman, and K. Lee. Imaging skin pathology with polarized light. *Journal of Biomedical Optics*, 7(3):278–523, 2002. 2

[14] F. Kahl. Multiple view geometry and the L$\infty$ norm. In *International Conference on Computer Vision*, volume 2, pages 1002–1009, 2005. 5, 6

[15] N. Kollias. *Bioengineering of the Skin*, chapter 7, pages 95–104. CRC Press, 1997. 2

[16] N. Kollias and G. N. Stamatas. Optical non-invasive approaches for diagnosis of skin diseases. *Journal of Investigative Dermatology*, (7):64–75, 2002. 2

[17] D. G. Lowe. Distinctive image features from scale invariant keypoints. *International Journal on Computer Vision*, 60(2):91–100, 2004. 4

[18] I. Matthews and S. Baker. Active appearance models revisited. *International Journal on Computer Vision*, 60(2):135–164, 2004. 1

[19] S. Nayar, X. Fang, and T. Boult. Removal of specularities using color and polarization. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 583–590, 1993. 2

[20] J.-S. Pierrard and T. Vetter. Skin detail analysis for face recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007. 1

[21] G. N. Stamatas, M. Southall, and N. Kollias. In vivo monitoring of cutaneous edema using spectral imaging in the visible and near infrared. *Journal of Investigative Dermatology*, 2006. 3

[22] R. C. N. Studinski and I. A. Vitkin. Methodology for examining polarized light interactions with tissues and tissue-like media in the exact backscattering direction. *Journal of Biomedical Optics*, 5(3):330–337, 2000. 2

[23] A. N. Yaroslavsky, V. Neel, and R. R. Anderson. Demarcation of nonmelanoma skin cancer margins in thick excisions using multispectral polarized light imaging. *Journal of Investigative Dermatology*, 2003. 3