Ridge Matching Based on Maximal Correlation in Transform Space

Mingzhi Tian[1], Jihang Wang[1], Samantha Horvath[2], John Galeotti[2], Vijay Gorantla[3], George Stetten[1,2]

[1]*Department of Bioengineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA, USA*
[2]*The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA*
[3]*University of Pittsburgh Medical Center, Pittsburgh, PA, USA*

Abstract

Image matching, a common technique in Computer Vision to identify objects, persons, locations, etc., is widely used in both military and civilian applications. For common image matching algorithms, results may vary when the raw images are captured under different lighting conditions. To reduce the unwanted influence from ambient lighting, we propose a novel method to match images that contain ridge features. The new method uses an established ridge detection algorithm to reduce raw images to sets of ridge points, each point defined by its orientation and location. To perform ridge matching, we find the pair-wise transform between every ridge point from one image and every ridge point from another. The result is a point cloud in transform space. The correlation between two sets of ridge point is equivalent to the density of the point cloud, computed by convolving the point cloud with a blurring kernel. The best match is found as the location in transform space at which the correlation reaches global maximum. We tested the new method on two image pairs; the first image pair contained artificial ridge features and the second pair was sampled from a high resolution image of the human palm. Both tests returned accurate results.

Keywords: Image Matching, Ambient Lighting, Ridge Point, Transform Space

Introduction

The Ambient Lighting Problem

Image-based positioning systems are used for advanced medical imaging techniques. When a set of 2D images—such as acquired by a conventional ultrasound scanner—is used to construct a 3D volume of a structure inside the human body, we require the continual position of the scanner relative to the patient. With an *a priori* 3D image map of the exterior surface of the patient reconstructed from multiple high resolution images, subsequent images from a small mobile camera can be matched with a correct projection rendered from the 3D map to compute the position of the camera relative to the patient. When the small mobile camera is mounted on the ultrasound scanner, the relative position of the scanner thus determined can be used to construct a 3D volume [1]. A significant problem with this camera-based approach is that images are affected by ambient lighting conditions. The performance of common image matching algorithms is penalized when variations in ambient lighting affects the intensity of image pixels, causing the local features used in matching to become unstable.

Motivation for Ridge Matching

The ridges on the skin are inherent physical structures, often represented by connected groups of salient ridge points within an image. Under normal lighting conditions, these features are resistant to changes in light source location, light source intensity, and shadow patterns. Another key advantage of ridge features is their inherent orientation property. Every ridge point has a well-defined ridge direction, tangent to the ridgeline at the given ridge point. The ridge direction provides additional constraints for computing the rotation between different images. Since ridge features are common on the surface of the human body, the same ridge structure is likely to be detected in images that have significant overlap. By using ridge features, a fast and closed form solution can be reached for matching two camera images with a rigid transform [3] [4].

Methods

Representation of Ridge Features

Our new method first extracts the location of ridge points in the images, using preprocessing algorithm based on an established scale-invariant ridge detection method [2]. The algorithm takes a raw grayscale image **I** as the input and returns a black and white image **BW**

$$\mathbf{H}(x_0, y_0) = \begin{bmatrix} \dfrac{\delta^2 \mathbf{I}}{\delta x^2} & \dfrac{\delta^2 \mathbf{I}}{\partial x \partial y} \\[2mm] \dfrac{\delta^2 \mathbf{I}}{\partial x \partial y} & \dfrac{\delta^2 \mathbf{I}}{\delta y^2} \end{bmatrix}_{(x_0, y_0)} = \begin{bmatrix} \cos\theta_0 & -\sin\theta_0 \\ \sin\theta_0 & \cos\theta_0 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos\theta_0 & \sin\theta_0 \\ -\sin\theta_0 & \cos\theta_0 \end{bmatrix} \quad \text{(Eq. 1)}$$

$$\mathbf{S} = \begin{bmatrix} x_1 & y_1 & \theta_1 \\ x_2 & y_2 & \theta_2 \\ \vdots & \vdots & \vdots \\ x_n & y_n & \theta_n \end{bmatrix} \quad s.t \ \mathbf{BW}(x_i, y_i) = 1 \quad \text{(Eq. 2)}$$

$$t(v_1, v_2) = \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta\theta \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ \theta_1 \end{bmatrix} - \begin{bmatrix} \cos\Delta\theta & -\sin\Delta\theta & 0 \\ \sin\Delta\theta & \cos\Delta\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ \theta_2 \end{bmatrix} \quad \text{(Eq. 3)}$$

as the output. A pixel at the image location (x, y) is considered a ridge point if $\mathbf{BW}(x, y) = 1$.

Each ridge point has an inherent orientation, which is the local direction along the ridge line. To compute the orientation at the ridge point (x_0, y_0), we consider the local Hessian matrix $\mathbf{H}(x_0, y_0)$. Since $\mathbf{H}(x_0, y_0)$ is a symmetric real valued matrix, we may perform eigenvalue decomposition on $\mathbf{H}(x_0, y_0)$ to obtain an orthonormal set of eigenvectors (Eq. 1).

Without loss of generality, let $|\lambda_1| < |\lambda_2|$. Then the unit eigenvector corresponding to λ_1 represents the principle direction of least curvature. The reason to find this eigenvector is that, in a ridge-like structure, the principle direction of least curvature points along the ridge line. From Eq. 1, the eigenvector that corresponds to λ_1 is the unit vector $[\cos\theta_0 \ \sin\theta_0]^T$, which can be described by a scalar angle θ_0 that ranges from -180° to 180°. The orientation of the ridge point at (x_0, y_0) is therefore mathematically defined as the scalar angle θ_0, for which $\mathbf{H}(x_0, y_0)[\cos\theta_0 \ \sin\theta_0]^T = \lambda_1[\cos\theta_0 \ \sin\theta_0]^T$.

The original image $\mathbf{I}$ is reduced to a set of ridge points, represented by a matrix $\mathbf{S}$ defined in Eq. 2.

Each row of $\mathbf{S}$ contains the necessary parameters to describe a ridge point: the ridge point location (x_i, y_i) in the image, and the ridge point orientation θ_i. The subscript n in Eq. 2 represents the total number of ridge points found in the original image, and the value is n is typically less than 5% of the total number of pixels in $\mathbf{I}$. The resulting matrix $\mathbf{S}$ is used as input to a ridge matching algorithm, described next.

Operations in Transform Space

To match two images $\mathbf{I}_1$, $\mathbf{I}_2$ with a rigid transform, we need to find the best overall translation and rotation between the two images.

We define the Transform Space $\mathbf{K}$ as the set of all possible rigid transforms $\{\Delta x, \Delta y, \Delta\theta\}$ between the two images; the best match is represented by a location in $\mathbf{K}$. Since the images have been reduced to two ridge feature matrices $\mathbf{S}_1$ and $\mathbf{S}_2$, each pair of ridge points, $v_1 = (x_1, y_1, \theta_1) \in \mathbf{S}_1$, $v_2 = (x_2, y_2, \theta_2) \in \mathbf{S}_2$, are correlated by a rigid transform $t: \mathbf{S}_1 \times \mathbf{S}_2 \rightarrow \mathbf{K}$ given by (See Eq. 3 above)

As a consequence of Eq. 3, if $\mathbf{I}_1$ and $\mathbf{I}_2$ have their top and left boundaries aligned (allowing for different sized images with their origins in the upper left corner), by applying the rigid transform t to image $\mathbf{I}_1$, the locations and orientations of v_1 and v_2 will coincide.

To find the best rigid transform between $\mathbf{I}_1$ and $\mathbf{I}_2$, we map every pair of ridge points $v_1 = (x_1, y_1, \theta_1) \in \mathbf{S}_1$, $v_2 = (x_2, y_2, \theta_2) \in \mathbf{S}_2$ to a vector in the Transform Space, and the result is a cloud of points in Transform Space $\mathbf{K}$. Every point in the resulting point cloud represents a potential transform mapped from a unique pair of ridge points. When a large number of ridge point pairs map to the same transform, that particular transform is more likely to be correct. In an intuitive interpretation, if the best rigid transform is applied to $\mathbf{I}_1$, the highest number of ridge points will approximately coincide.

$$D(\Delta x, \Delta y, \Delta \theta) =$$

$$\left[\sum_{v_i \in S_1} \sum_{v_j \in S_2} \delta\left(\Delta x - \Delta x(v_i, v_j)\right) \delta\left(\Delta y - y(v_i, v_j)\right) \delta\left(\Delta \theta - \Delta \theta(v_i, v_j)\right) \right] * f(\Delta x, \Delta y, \Delta \theta) \quad (4)$$

Maximal Correlation

In practice, no two ridge point pairs will map to the exact same transform when the images acquired with noise and distortion are sampled at a finite resolution. However, when a good match exists between two images, the point cloud tends to form a dense cluster near an optimum location in the Transform Space **K**. A measurement of point cloud density at every location in **K** provides us the correlation between S_1 and S_2 for every possible rigid transform. To do so we define a density function D: $\mathbf{K} \rightarrow \mathbf{R}$, where

In Eq. 4, each vector in Transform Space **K** is treated as an impulse function of three variables. The point cloud in **K** is therefore represented as a finite sum of impulse functions. We compute density function by convolving the point cloud and a blurring kernel f. Here, the kernel f has the value 1 in the cuboidal region of size $1 \times 1 \times 0.2$ centered at $(0, 0, 0)$ and has the value 0 elsewhere.

The density function D reaches a global maximum at the optimal transform. The best match is found as the location in Transform Space at which assumes the maximal value.
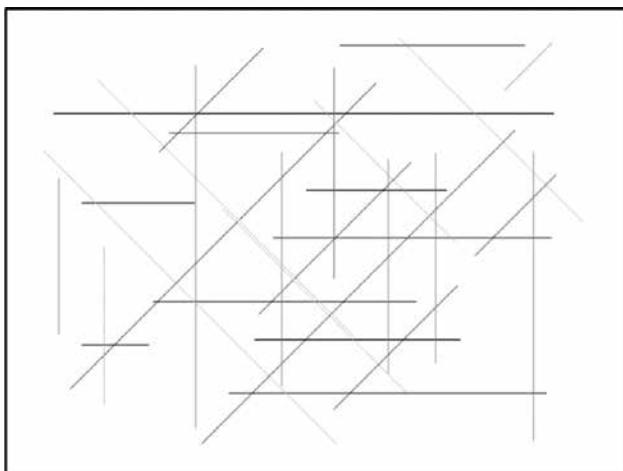
Results

Test on Artificial Ridge Features

We first tested our method on artificially constructed ridge features. Figure 1 shows a large image containing these artificial features. Every ridge point in this particular set has only four possible orientations: θ = -45°, 0°, 45°, or 90° The location and orientation of every ridge point are precisely known and serve as ground truth for validating our ridge matching algorithm.

Two small patches (See Figure 2) were selected from the large artificial ridge map to simulate the binary images $\mathbf{BW_1}$, $\mathbf{BW_2}$ after ridge extraction. The offsets were $(0, 0)$ and the rotation between them was 50°. The patches were used to generate the point cloud in Figure 3. The density of the point cloud was calculated and displayed separately in Figure 4 and Figure 5. The optimal transform was found as $(0, 1, 50°)$, accurate to a single pixel.

Test on Real Sampled Images

The ridge matching algorithm is next tested on real images of the human palm. Two patches (See Figure 6) were sampled from a larger palm image of the palm. The offset between the sampled patches was $(80, -20)$ and the rotation between the patches was 80°. Figure 7 shows the binary images produced by the ridge detection algorithm. Figure 8 shows the point cloud generated from the ridge feature matrices. Figure 9 and Figure 10 display the density map of the point cloud in Figure 8. The maximal correlation density
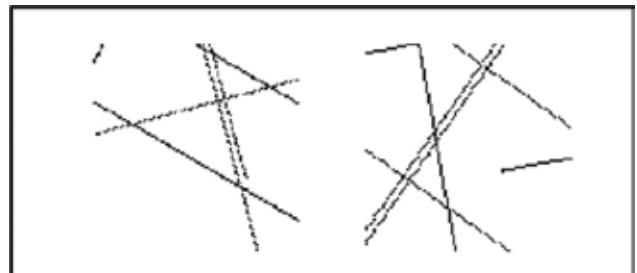

Figure 1. Image containing artificial ridge features

Figure 2. Binary images selected from the large artificial ridge map
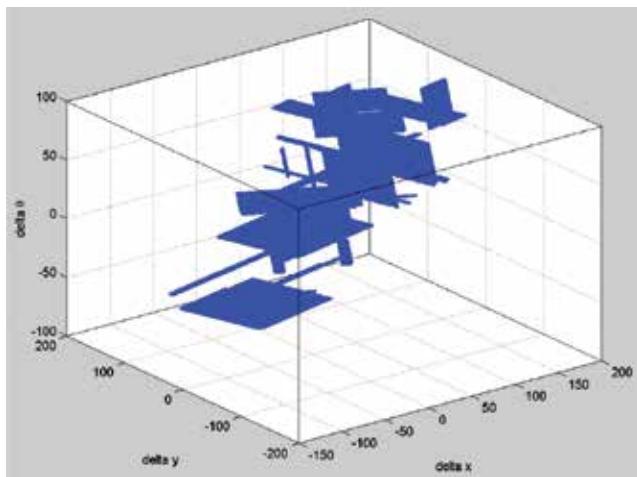
Figure 3. Point cloud in Transform Space generated from two sets of artificial ridge features
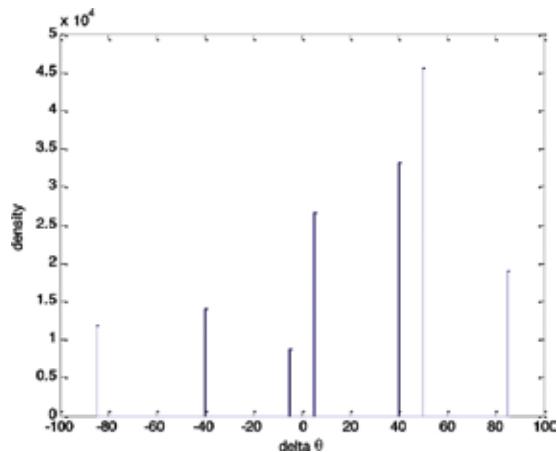

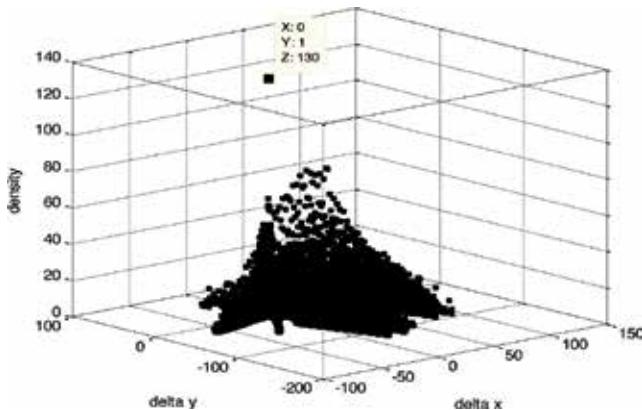**Figure 4.** Projection of the density function onto the $\Delta\theta$ axis

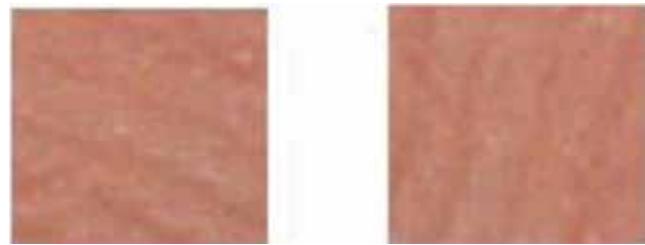**Figure 5.** Cross section of the density function at $\Delta\theta = 50°$

Figure 6. Sampled images of the human palm

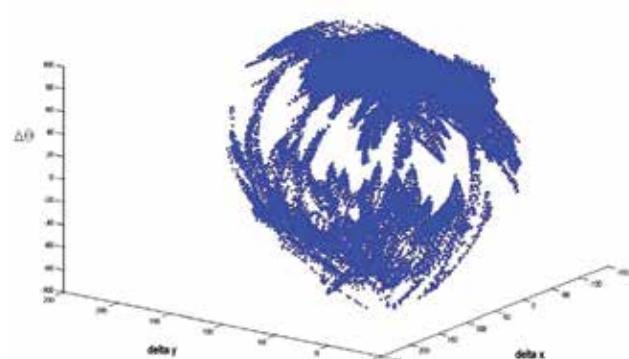**Figure 7.** Binary images containing ridge points detected in the palm images

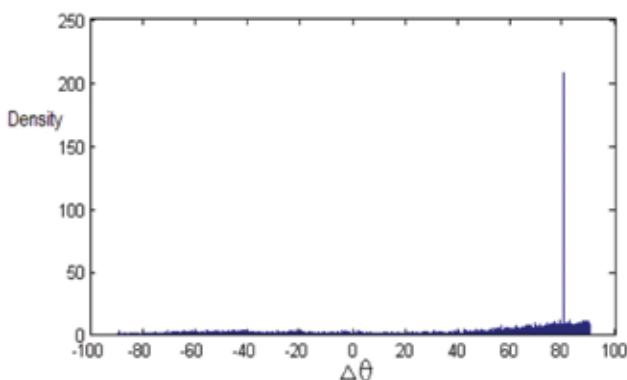**Figure 8.** Point cloud generated from real ridge sets

Figure 9. Projection of the density function onto the Δθ axis

Figure 10. Cross section of the density function at Δθ = 50°

occurs at (80, -19, 80°), also accurate to a single pixel, reasonable given sampling error.

Discussion

The artificial ridge features produced strong outlier clusters in the Transform Space. This effect can be attributed to how the artificial ridges were constructed, where all ridge points were forced to take one out of four possible orientations and all ridge lines are long continuous straight lines. There is a risk that when the images being matched contain a large number of repetitive ridge features, the correct match will reliably result from our method. However, in our particular experimental case, the artificial ridge test demonstrated that our method is capable for matching images to within a single pixel of error, warranting further test on real image datasets.

The real palm images produced equally accurate results. The point cloud, compared to that generated from artificial features, is more scattered in Transform Space, but the global peak in the density function is also more prominent for real image datasets.

Conclusion

The strong response demonstrated to inherent ridge features allows operations without specialized markers on the skin. As the ridge feature possess inherent orientations, these features are likely to stay stable under normal variations in ambient lighting. Given the present challenge of matching two 2D images with a ridge transform, a process with 3 degrees of freedom (2 translations and 1 rotation), the fact that the individual ridge features also have the same equivalent
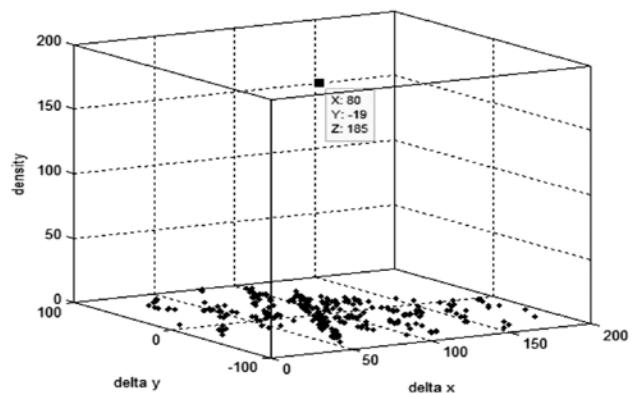
3 parameters (2 locations and 1 orientation) results in a closed form solution that is both fast and reliable. The accuracy of the matching algorithm requires that the images being matched contain ridge-like features. By modifying the preprocessing algorithm to detect other types of features such as edges or elliptical blobs, our proposed method can also be generalized to other types of feature matching. We are also working to include additional degrees of freedom, such as scale changes between two images, which can be accommodated by using higher dimensional votes in transform space than simple points.

References

[1] Galeotti et al, Image Guided Therapy Workshop, Oct 2011.

[2] Lindeburg, International Journal of Computer Vision 30(2), 79-116 (1998)

[3] Li, Yuhai, et al. "A fast rotated template matching based on point feature." Proceedings of the SPIE 6043 (2005): 453-459. MIPPR 2005: SAR and Multispectral Image Processing.

[4] Wen-Chia Lee. "A Fast Template Matching Method for Rotation Invariance Using Two-Stage Proces." Sept, 2009. IEEE, ISBN: 978-1-4244-4717-6.