Projected Texture for Hand Geometry based Authentication

Avinash Sharma Nishant Shobhit Anoop Namboodiri
Center for Visual Information Technology
International Institute of Information Technology, Hyderabad, INDIA
{avinash@research.,nishant@students.,anoop@}iiit.ac.in

Abstract

We propose a novel approach to 3D hand geometry based person authentication using projected light patterns. Instead of explicitly computing a depth map of the palm for recognition, we capture the depth information in the deformations of a projected texture pattern, and use it directly for recognition. The deformed pattern is characterized using local texture measures, which can encode the certain depth characteristics of the palm. An authentication system built using the proposed technique achieves an equal error rate of 0.84% *on a dataset of* 1341 *samples collected from* 149 *users, as opposed to* 4.03% *using traditional 2D features on an identical dataset. The approach is robust as well as computationally efficient and could be applied to other 3D object recognition problems as well.*

1. Introduction

Recognition of 3D objects from their images is a challenging problem as the depth information of the scene is lost during image acquisition. The loss of depth information is critical for certain applications such as 3D hand-geometry based authentication, where subtle variations in the 3D shape provide important clues about the identity of the subject. The traditional approach to overcome this problem for any 3D biometric authentication has been to explicitly compute the depth information using dedicated range scanners or computationally intensive algorithms using multi-view geometry. A shape descriptor is then computed from the recovered 3D shape, which is used for classification. The process of computing depth is often sensitive to accurate determination of point correspondences, which could be challenging when prominent features are not present in the object. This will in turn affect the reliability of the shape descriptors for classification.

The approaches to object recognition can be classified into two categories based on the nature of features that are extracted from the image. The first class of approaches uses object-related features such as 3D or 2D shape for recognition. This involves segmentation of the object and inferring properties of the object that were lost during imaging. The second class of approaches uses properties of images of objects such as intensity variations, color, texture, etc. Specifically the use of texture has proven to be extremely reliable and accurate for the purpose of classification and has hence been used for biometric modalities such as iris, fingerprints, palm prints, etc.

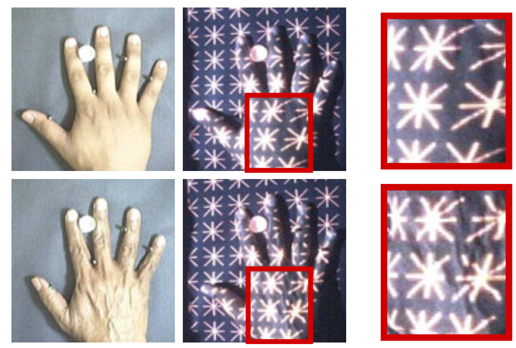

Figure 1. Deformations in projected texture.

In this light, we propose to derive features of the object that are dependent on the depth, thus characterizing the 3D shape of the object well; while avoiding explicit depth computation. The primary idea is to use structured lighting to illuminate the object during the acquisition phase. The deformations in the projected pattern are dependent on the depth variations in the scene, and hence the deformed pattern can be thought of as encoding the depth information within it. By this process, we have transformed the ill-posed and computationally intensive problem of deriving object shape based features, into the well understood and efficient domain of characterizing the intensity variations within the image. Figure 1 shows two hand images whose depth maps are very similar. However, the variations in the depth introduce significant changes in the deformation of the projected texture, which can be readily used for recognition.

Textures, which are deterministically or stochastically repeating patterns of intensities in an image, have been studied extensively for computer vision applications. Textures on object surfaces have traditionally been used for tasks such as segmentation, image retrieval, and biometric authentication, among others. In the context of recognition, certain objects with distinguishable shape characteristics might not have prominent or discriminative textures that can tell similar classes apart. This shortcoming can be overcome by projecting structured illumination onto them. The deformations in the projected texture/pattern, treated as another texture, not only give us an estimate of the 2-D geometry of the object but also of the depth variations along its surface. Wavelet methods of texture analysis, such as Gabor filter banks, capture these deformations without having to compute point correspondences and handling occlusions.

1.1. Hand Geometry based Authentication

Measurements of the human palm, such as the length and width of fingers and the 3D palm profile are known to contain some amount of identity information. Sildauskas [10] patented the first electronic hand geometry based identity verification apparatus, and several commercial systems have been developed since then. Jain *et al* [6] outlined the challenges in such an authentication system and proposes a simple set of hand measurements, inspired by the previous work. Even the most recent hand geometry algorithms [4] use extentions of the set of features outlined in [6]. The research in 2D hand geometry based authentication has progressed primarily in three different directions:

The first set of algorithms tried to include additional measurements of palms such as area, perimeter, distances between specific feature points on the palm, etc. [4] to improve the verification accuracy. Even though the results showed improvements on the prior art, the comparisons are limited. A second direction was to integrate hand geometry along with other biometric traits to achieve high recognition performance. Fingerprints [11] and Palm prints [7, 12] are ideal candidates for this due to their ease of acquisition along with the hand geometry. A third set of algorithms look at generic techniques to improve the classification process used for verification, such as feature discretization [8], use of error correcting codes and more powerful classifiers [4], etc.

The use of 3D information in hand-geometry based authentication is limited to adding partial depth information computed from the profile view, usually captured using a slanting mirror. The use of depth information of the hand has the potential to improve the recognition and verification performance of hand geometry based systems. The most promising approach for use with hand geometry based authentication is the use of structured lighting [13, 9], due to its robustness and the availability of a controlled environment during imaging. Cofer and Hamza [1] proposed the use of dot patterns to compute correspondences, and hence the depth, at specific points on the palm. The recovered depth is used along with silhouette features for recognition. Faulkner [3] proposed the use of light stripes instead of dot pattern to compute the correspondences. However, both of the above approaches aim to recover partial depth information, which in turn is used along with 2D object features for authentication.

In this paper, we propose the use of a projection pattern on the palm, which gets deformed according to the depth variations of the palm. The texture measure computed from local sub-windows of the captured image can be used to characterize the hand geometry. Texture measures inherent in a biometric traits such as palm prints [7], fingerprints [5] and iris patterns [2] have been used extensively for identity verification. We propose a similar approach for hand geometry based authentication on projected texture. The primary contribution of the present work is that we are able to map the problem of computation of 3D features for recognition into that of object recognition using a texture description. To the best of our knowledge, this is the first attempt at characterizing the shape of an object using deformed projected textures for the purpose of recognition.

2. Projected Texture for Recognition

The key idea of the approach, as described before, is to encode the depth variations in an object as deformations of a projected texture. There are primarily two categories of objects that we might want to characterize. The first class of objects, such as manufactured parts and human palm, are characterized by their exact 3D shape, while the second class of objects are characterized by the stochastic variations in depth such as 3D textured surfaces. Although the proposed approach can be adapted for use in either scenario, we concentrate on the first class of objects in this work.

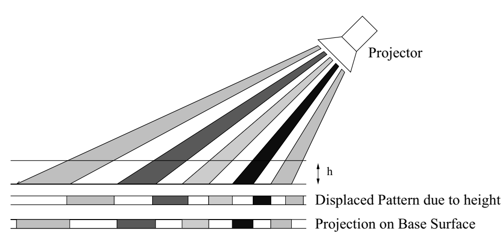

Figure 2. Shift in captured pattern due to object height.

The object to be recognized is placed under controlled pose and a specific texture pattern is projected on it. The projected pattern (or the original texture), falling on the surface containing the object, gets modified according to the depth map of the object. These transformations can be primarily classified into two categories:

- *Pattern Shift*: The position where a particular projected pattern is imaged by the camera depends on the abso-

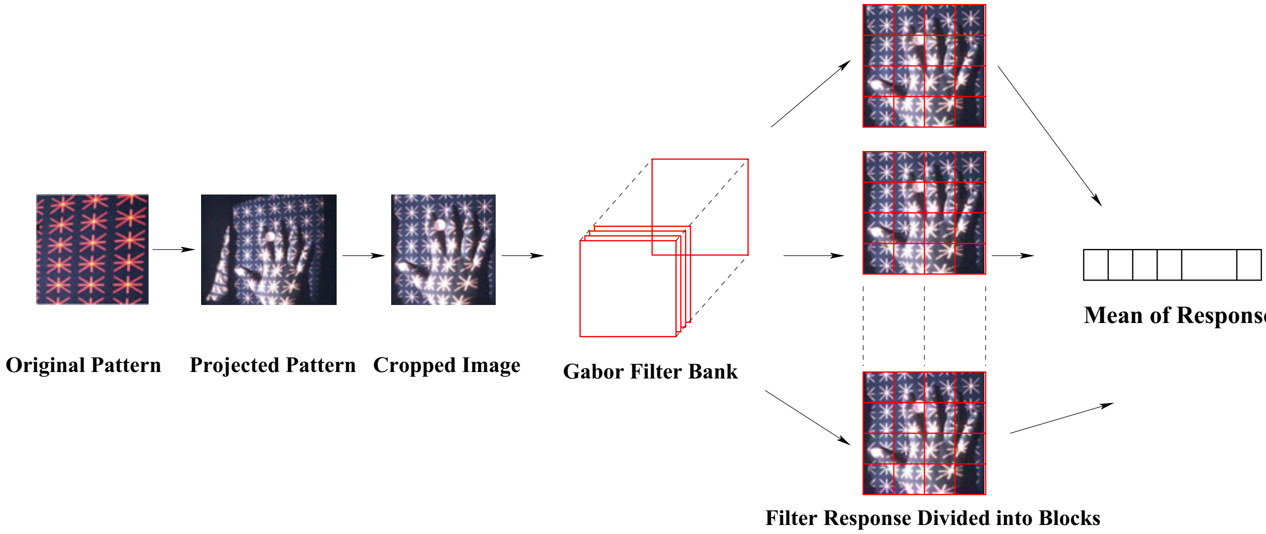

Figure 3. Computation of the proposed projected texture based features.

lute height from which the pattern in reflected by the object. Figure 2 illustrates this with a cross section of a projection setup. Note that the amount of shift depends on the height difference between the objects as well as the angle between the projector axis and the plane of the palm.

- *Pattern Deformation*: Any pattern that is projected on an uneven surface gets deformed in the captured image depending on the change in depth of the surface (see Figure 1). These deformations depend on the absolute angle between the projector axis and the normal to the surface at a point as well as its derivative.

We claimed that the deformations of the projected texture contain sufficient information to characterize the object. To understand the process, we inspect a specific window of the image with a particular pattern in the middle. When the height of the object within the window changes, the pattern either shifts away from the window or the lines in the pattern changes direction, as we noted above. In both cases, the local frequency characteristics of the window changes considerably. We capture this by the responses of a bank of Gabor wavelets in the window. To encode the shape of the object, we partition the captured image into smaller windows, which are then represented using the mean filter responses of Gabor wavelets in various directions and scales.

One should note that the amount and nature of deformations depend on the relative positions of the projector, the camera and the object being imaged, as well as the nature of pattern being projected. That is, even minor variations in the height of an object can cause large variations in the texture pattern. However, the pattern to be projected should be carefully chosen for an application, so that the variations within the sub-windows are significant.

2.1. Designing a Projection Pattern

The choice of an appropriate projection pattern is important due to a variety of factors:

- For the deformation to be visible at any point in the captured image, the gradient of the texture at that point should not be zero in the direction of gradient of the object depth.
- The location and direction of the texture pattern should induce a discernable response in the filter set used to characterize the texture.
- The density of the projected pattern or its spatial frequency should correspond to the frequency of height variations to be captured. Hence, analyzing the geometry of an object with a high level of detail will require a finer texture, whereas in the case of an object with smooth structural variations, a sparse one will serve the purpose.
- Factors such as the color, and reflectance of the object surface should be considered in selecting the color, intensity and contrast of the texture so that one can identify the deformations in the captured image.

For the purpose of hand geometry based authentication, we have selected a repetitive star patten that has gradients in four different directions. This will allow us to capture depth variations in different directions within a window. The width of the lines and the density of patterns in the texture were selected experimentally so that it captures the height variations between the palms at the angle of projection selected.

2.2. Characterizing Hand Geometry

Once the pattern is selected, we need to characterize the deformations induced by the height variations of the object. For hand-geometry based verification, we divided the image into 64 sub-windows (a 8×8 grid). Each sub-window is then characterized by the responses of Gabor filters that captures the local frequencies and their orientations.

A Gabor function is a Gaussian function, modulated by a complex sinusoid. A simplified form of the filter, $G(x, y)$, may be written as:

$$\begin{aligned} G_{\sigma,\phi,\theta}(x, y) &= g_\sigma(x, y) \cdot e^{2\pi j\phi(x\cos\theta + y\sin\theta)} \\ g_\sigma(x, y) &= \frac{1}{\sqrt{(2\pi)}\sigma} e^{-(x^2+y^2)/2\sigma^2}, \end{aligned}$$

where θ is the orientation of the sinusoid with frequency ϕ, and $g_\sigma(x, y)$ is a Gaussian with scale parameter σ. In our experiments, we use a bank of 24 filters with 8 orientations ($\theta = 0, \pi/8, 2\pi/8, \cdots, 7\pi/8$) and 3 radial frequencies, controlled by the frequency of the sinusoid. The feature vector representing a sample image is computed as follows (see Figure 3).

The image is converted to gray scale and the area of the image that contains the palm is cropped. The pixel values are then normalized to have a specific mean and variance for the image. The resultant image is then convolved with each of the 24 Gabor filters and the mean of the filter responses are computed for each sub-window. In our experiments, we divided each response image into 64 sub-windows (8×8). This resulted in a feature vector of dimension 1536.

3. Experimental Results and Analysis

To analyze the discriminative power of the projected texture based features, we check the verification performance on a dataset of 1341 hand images collected from 149 subjects. Each subject provided 9 images with the projected texture and 9 with uniform illumination to serve as a comparison dataset for traditional 2D approaches.

The image capturing setup is similar to that discussed in Jain *et al* [6], where pegs are used to guide the placement of the palm. However, unlike the popular peg-based datasets, where the placement of the hand is controlled, we encouraged the users to vary the hand pose within the peg limits to make the dataset more realistic as in unsupervised scenarios. The surface of placement of the palm was darkened to facilitate the segmentation process for 2D image analysis. The illumination over the area of the palm could be either uniform or from a projector that is placed at an angle

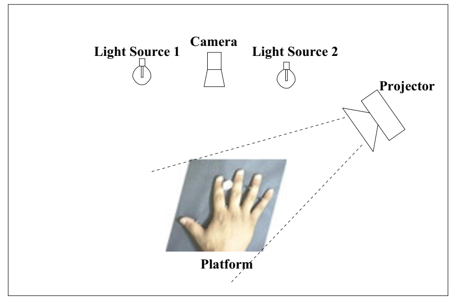

Figure 4. Imaging setup.

Figure 5. The hand under structured illumination as well as normal lighting. The mirror to the left enables us to compute height of palm at specific points.

to the palm (see Figure 4). The images are captured by a camera located directly above the palm with its optical axis perpendicular to the palm surface. A reflecting mirror fixed by the side of the imaging surface helps to capture a side view of the palm, and thus to include thickness of fingers as described in [6]. Each user provided 9 hand images with uniform illumination, as well as with projected texture (see Figure 5). The users were asked to remove their hand and replace it for each image captured, with limited variations in hand pose.

Due to the variations in hand pose, our dataset is much more challenging than popular ones using peg based approach (see Figure 6). Note that users introduced considerable variations in the pose, even when limited by the pegs. Similar variations were introduced for the projected texture dataset also. Due to these variations in pose, traditional approaches for 2D feature extraction that assumes peg-based imaging will fail in many samples. Hence we have tracked the finger locations and computed the features appropriately. We also verified and manually corrected any of the 2D features that were incorrectly computed due to pose variations.

3.1. Feature Extraction and Matching

For computing the 3D features, the images are cropped, converted to gray scale, and the pixel values are normalized to reduce lighting variations during imaging. For the purpose of comparison, we compute two different 2D feature sets from the samples with uniform illumination, in addition to those from the projected texture. The feature sets used are:

Figure 6. Samples from 3 users in our dataset: Note that the pose varies considerably, even with the use of pegs.

- *Feat-1*: The set of 17 feature proposed by Jain *et al* [6], computed from the width and length of fingers and palm as well as the height of the index finger computed from the image reflected on the mirror.

- *Feat-2*: A set of 10 features proposed by Faundez-Zanuy *et al* [4], including 5 finger lengths, area of the palm, the contour length and distance between specific points on the palm contour.

- *Feat-3*: The proposed projected texture based features, computed from filter responses from 64 sub-windows.

Comparisons with 3D hand geometry approaches such as Cofer and Hamza [1] and Faulkner [3] could not be carried out as the patents does not provide sufficient information about the exact nature of feature extraction. Moreover, as we mentioned before, our approach does not require depth computation or even segmentation of the palm, and hence is comparable to the image based approaches in spirit and complexity.

The primary aim of the experiments is to compare the proposed feature set to the traditional image based features. For this reason, we have avoided complex classifiers or post processing techniques as proposed in Faundez-Zanuy *et al* [4]. One of the best indicators of the discriminating power of a feature set is the ROC curve induced by distances computed in the corresponding feature space. The ROC curve indicates the level of separation between the genuine and imposter distance distributions. We used a simple Euclidean distance to compute the distance between feature vectors in all three cases. Note that the performance of the classifiers in each case could be improved by more complex classifiers or post processing techniques. Hence the accuracies reported here should be used only for comparison of the feature spaces, and not as an indicator of the absolute discrimination power of any of the feature sets.

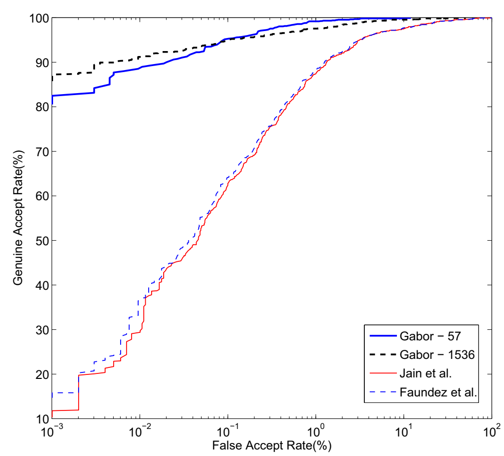

Figure 7. ROC curves for two 2D feature based, and the proposed projected texture based approaches.

Figure 7 gives the ROC curves obtained from the three feature sets mentioned above. The Equal Error Rate (EER), or the rate at which false rejects equals false acceptance rate, is a single indicator that can be computed from the ROC curve. The EERs for the *Feat-1* and *Feat-2* were 4.06% and 4.03% respectively, while the proposed feature set achieved and EER of 1.91%. Clearly the projected patterns induce a large amount of discriminating information into the computed features. In addition to the equal error rate, we note that the genuine acceptance rate continues to be above 80%, even at false acceptance rates of 0.001% for the proposed features, while the performance of the 2D image based features degrade considerably at this point.

We also conducted an experiment in feature selection to choose a subset of the 1536 features that would help us in reducing the computations required. We note that even with just 57 features out of 1536, the ROC curve is similar to that of the complete feature set. Moreover, the equal error rate improves to 0.84% with the reduced feature set. This is possible as the selection process avoids those sub-windows, where the intra-class variations in pose are high.

Figure 8. Selected feature windows

Figure 8 shows the windows corresponding to the most discriminative 12 features. Note that the features belong to windows that are at the edges of the fingers as well as on the palm surface, which indicates that the depth information of the palm is also used for authentication, in addition to the shape of the fingers. The presence of a window outside the palm region could be because it encodes the relative brightness of projected pattern, which in turn encodes the skin color.

Another interesting observation is that a weighted combination of the distance scores from 2D and texute-based features did not improve the performance. Evidently, the projected texture encodes most of the information that is contained in the 2D features, in addition to the 3D information of the hand surface.

4. Conclusions and Future Work

We have proposed a new approach to hand-geometry based authentication using the projection of a structured light pattern during image acquisition. We note that the computation of textural features from specific local windows can yield a feature vector that is far more discriminative than traditional 2D object features used for hand geometry based authentication. The approach is computationally efficient and the time taken is comparable to that of the 2D image based authentication. Moreover, the approach is robust to occlusions and noise as opposed to 3D hand geometry systems that need to explicitly compute a depth map of the hand.

However, a several issues still remain unaddressed in applying the recognition approach to generic objects. The method is sensitive to the relative positioning of the camera, the projector, and the object. Object reflectance and transparency might be another interesting area to explore. Our approach is extensible to 3D textures as well. Temporal variations in dynamic texture deformations could also give us a cue towards designing optimal classifiers for recognition.

References

[1] D. Cofer and R. Hamza. Method and apparatus for detecting objects using structured light patterns. *US Patent No. 7176440*, 2007. 2, 5

[2] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1148–1161, 1993. 2

[3] K. Faulkner. Apparatus and method for biometric identification. *US Patent No. 5335288*, 1994. 2, 5

[4] M. Faundez-Zanuy, D. A. Elizondo, M. Ferrer-Ballester, and C. M. Travieso-González. Authentication of individuals using hand geometry biometrics: A neural network approach. *Neural Processing Letters*, 26(3):201–216, 2007. 2, 5

[5] A. K. Jain, S. Prabhakar, L. Hong, and S. Pankanti. Filterbank-based fingerprint matching. *IEEE Transactions on Image Processing*, 9(5):846–859, 2000. 2

[6] A. K. Jain, A. Ross, and S. Pankanti. A prototype hand geometry-based verification system. In *Proceedings of the AVBPA'99*, pages 166–171, Washington D.C., Mar. 1999. 2, 4, 5

[7] A. Kumar and D. Zhang. Personal recognition using hand shape and texture. *IEEE Transactions on Image Processing*, 15(8):2454–2461, August 2006. 2

[8] A. Kumar and D. Zhang. Hand-geometry recognition using entropy-based discretization. *IEEE Transactions on Information Forensics and Security*, 2(2):181–187, June 2007. 2

[9] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. *CVPR03*, pages 195–202, 2003. 2

[10] D. P. Sidlauskas. 3d hand profile identification apparatus. *US Patent No. 4736203*, 1988. 2

[11] K.-A. Toh, W. Xiong, W.-Y. Yau, and X. Jiang. Combining fingerprint and hand-geometry verification decisions. In *Proceedings of Audio- and Video-Based Biometric Person Authentication*, volume 2688, pages 1059–, Springer Berlin, 2003. 2

[12] J. Wu and Z. Qiu. A hierarchical palmprint identification method using hand geometry and grayscale distribution features. In *Proceedings of International Conference on Pattern Recognition*, volume 4, pages 409–412, 2006. 2

[13] L. Zhang, B. Curless, and S. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. *3DPVT02*, pages 24–36, 2002. 2