Multi-Scale Azimuthal Projection Distance Image of Normal Maps for 3D Facial Skin Texture Analysis

Alassane Seck
als31@aber.ac.uk
Hannah Dee
hmd1@aber.ac.uk
Bernard Tiddman
bpt@aber.ac.uk

Computer Science Department
Aberystwyth University
Aberystwyth, UK

Abstract

In this paper we propose a multi-scale Azimuthal Projection Distance Image scheme for analysing 3D texture and use it in classifying wrinkles and pores on high resolution facial normal maps. Azimuthal Projection Distance Images (APDI) have been widely used for mapping normals' orientation to scalar values on a tangent plane. The multi-scale scheme we introduce computes APDIs at different resolutions of the normal map. We propose a geodesic-based method of scaling normal maps and modify the original formulation of the APDI which only take into consideration polar orientation. We finally apply the approach in classifying facial skin conditions (large pores, wrinkles). Local Binary Patterns feature are extracted from the multi-scale APDI pyramid and put into a Multi-Layer Perceptron. The results we achieve are promising even though the dataset used is limited.

1 Introduction

3D surface data have been widely used in Pattern Recognition for extracting 3D invariant features [1, 6, 15]. These have been mostly used as key-points for various computer vision tasks such as 3D Object Recognition, 3D Saliency Detection, Mesh Alignment etc. Recent advances in 3D surface recovery allow capture of surface details with great precision and resolution. This opens new opportunities in machine vision and pattern recognition. One of these is the possibility of analysing texture directly from dense 3D surface data.

3D surface texture analysis has not seen as much interest as 2D texture analysis, although there are some studies that address the problem. Smith *et al* proposed a method for computing a co-occurrence matrix for normal maps [12]. Sandbach *et al* computed Local Binary Patterns on depth maps and Azimuthal Projection Distance Images of normal maps to classify 3D facial action units [11]. Peyre and Mallat proposed an interesting bandelets approach for compressing 3D surface geometry [9]. Other studies have used 3D patterns in analysing skin conditions. Warr *et al* demonstrated the considerable value added by augmenting the

© 2013. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

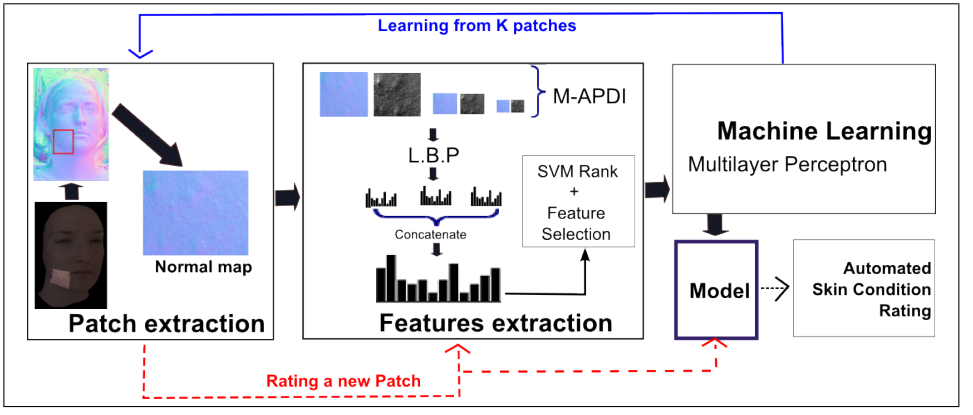


Figure 1: Algorithm overview

classical 2D ABCD (Asymmetry, Border, Color and Diameter) features with 3D pattern analysis in classifying benign and malignant Melanomas [2]. Koh *et al* used a 3D imaging system to quantify skin surface roughness and acne. Warr *et al* used first and second order differential forms of skin surface relief to describe skin lesion disruption.

In this paper, we investigate a multi-resolution Azimuthal Projection Distance Image scheme for analysing texture from high resolution facial normal maps. We propose a geodesic-based method for scaling normal maps and modify the original formulation of the Azimuthal Projection Distance Image in order to take into account the azimuthal orientation of the normals. We finally test the proposed approach by classifying skin conditions such as wrinkles and large pores from highly detailed facial normal maps.

In the first section, we recall the notion of Normal Map, give an introduction of Riemannian geometry on Normal Map and introduce our reformulation of the Azimuthal Projection Distance Image. In the second section present our Multi-resolution APDI algorithm. And finally we give in the last section our experimental results.

2 Background

In this paper we investigate methods for analysing 3D texture from high resolution normal maps. Even though normal maps are represented with a 2D RGB image, classical linear 2D texture analysis methods can not be applied directly. Each pixel of a normal map represents a normal's orientation and therefore they do not belong to a linear space. Summing two normals does not result in another normal as the result will not be a unit vector.

The approach considered in this paper is to project the normals onto a tangent Euclidean plane where linear texture analysis methods can be applied. We use Riemannian differential geometry elements to formulate our geodesic-based normal map rescaling. In this section, we present the concept of normal map, provide an overview of some Riemannian geome-

try elements, give a short introduction to Azimuthal Projection Distance Image and finally present our modification of its original formulation.

2.1 Normal Map

A normal map, also known as a needle-map, is an RGB image that represents surface's orientations by storing, in each of its pixel channels, a Cartesian component of a surface normal. Typically, the Red channel contains the X-component, the Green channel contains the Y-component and the Blue channel contains the Z-component. Generally the three channels are coded in the RGB image with a range of [0;255] and in such cases, they should be rescaled into the range [-1;1] before further usage. Figure 2 gives an example of a facial normal-map from the ICT3DRFE dataset [13]; the RGB visualisation gives an indication of how the different normals are oriented.

Figure 2: Example of a facial normal-map [13]

Normal maps are generally produced by photometric stereo techniques. They are widely used in computer graphics to enhance the level of detail of a low resolution geometry mesh, either by integrating them and producing a higher resolution mesh [8] [7] or by using bump-mapping techniques at rendering time [7]. Normal maps can be expressed in the object-space (which is the same space where the mesh geometry is expressed) or in the tangent-space. Generally, in the later case, a transformation of either the mesh or the normal-map between the two spaces is necessary before integrating them.

In the rest of the paper, we will define a normal-map as:

$$n(i,j)_{0\leq i\leq W}^{0\leq j\leq H}$$

With W and H respectively the normal-map width and the height. $n(i,j)$ will denote the normal value of the pixel (i,j) .

Riemannian Geometry on Normal Maps

A Riemannian manifold M is a space equipped, at each point $\mu \in M$, with a tangent bundle $T\mu M$ where a differentiable inner product is defined for any vectors $x, y \in M$ [14]. Two functions are defined for projections between the manifold and tangent spaces:

- The Logarithm function maps points from M to $T\mu M$
- The Exponential function maps points form $T\mu M$ to M

We assume that the normals lie in a Riemannian manifold. This allow us to use the Log_μ function to project them onto the tangent plane at μ, where linear operations can be performed, and Exp_μ to project them back in the normal space (Figure 2.1). In this paper we chose $\mu = z$ (i.e. a uniform tangent parallel to the image plane).

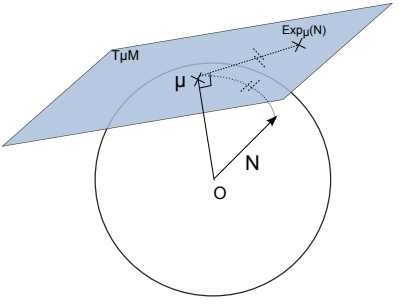


Figure 3: Logarithm and Exponential Chart on the Unit Sphere

2.2 Azimuthal Projection Distance Image

The Azimuthal Projection Distance Image is an extension of Azimuthal Equidistant Projection (used in Geography and Earth Science) to 3D image processing. It aims to project normals onto points in an Euclidean plane. The projection image is generated by taking each pixel as the absolute distance of the projected point from the centre of projection [11]. For the rest of this paper we will use the acronym "APDI" to refer to Azimuthal Projection Distance Image.

Formally, if $n(i,j)_{0\leq i\leq W}^{0\leq j\leq H}$ denotes a normal-map, the normal $n(i,j) = (n_x, n_y, n_z)$ at the pixel (i,j) is projected onto the point $n'(i,j) = (x', y')$ such as:

$$x' = k\cos\theta\sin(\phi - \phi_0) \tag{1}$$

$$y' = k(\cos\theta_0\sin\phi - \sin\theta_0\cos\theta\cos(\phi - \phi_0)) \tag{2}$$

With: $\theta = \frac{\pi}{2} - \arccos(n_z)$ and $\phi = \arctan(\frac{n_y}{n_x})$

The plane of projection is chosen to be parallel to the image plane, so for each normal, the centre of projection is such that: $\theta_0 = \frac{pi}{2}$ and $\phi_0 = 0$. So equations 3 and 4 become:

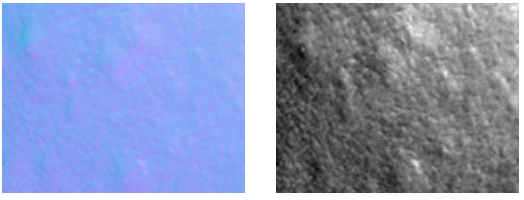
$$x' = k\cos\theta(i,j)\sin\phi(i,j) \tag{3}$$

$$y' = k\cos\theta(i,j)\cos\phi(i,j) \tag{4}$$

Finally, the Azimuthal Projection Distance Image is generated by creating an image with the same dimensions as the normal map and setting each pixel value to the absolute distance between the centre of projection and the projected point:

$$I(i,j) = \sqrt{x'^2 + y'^2} \tag{5}$$

Figure 4 shows an acned skin patch normal map and its Azimuthal Projection Distance Image.



(a) (b)

Figure 4: Normal-map (a) and APDI (b) of an acned skin patch

2.3 APDI Reformulation

In [□], Azimuthal Projection Distance Image is computed from normal-map by setting each pixel as the absolute distance of the projected point from the centre of projection (Equation 5). The problem with that formulation is that with a constant projection plane (parallel to the image plane), all the normals with the same polar angles will result in a same pixel value, even if they have different azimuthal angles. This means that the APDI does not encode the azimuthal orientation of the normals. To overcome this, we propose, instead of taking the absolute distance as pixel value, to take the arc formed by the projected point from the X'-axis on the projection plane (Figure 5).

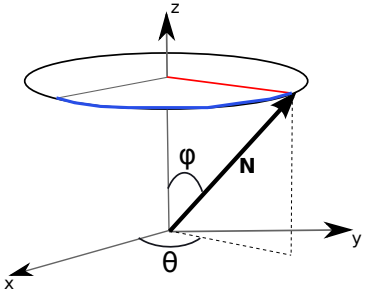


Figure 5: The distance from the centre of projection (red) is the same for all the normals with the same polar angle, while the arc (blue) from the X-axis varies with both the azimuthal and polar angles.

Thus, instead of equation 5, we propose the pixel values to be given by:

$$I(i,j) = \arctan 2(y',x')\sqrt{x'^2+y^2} \tag{6}$$

We give in figure 6 a comparison of the APDI and our modified version.

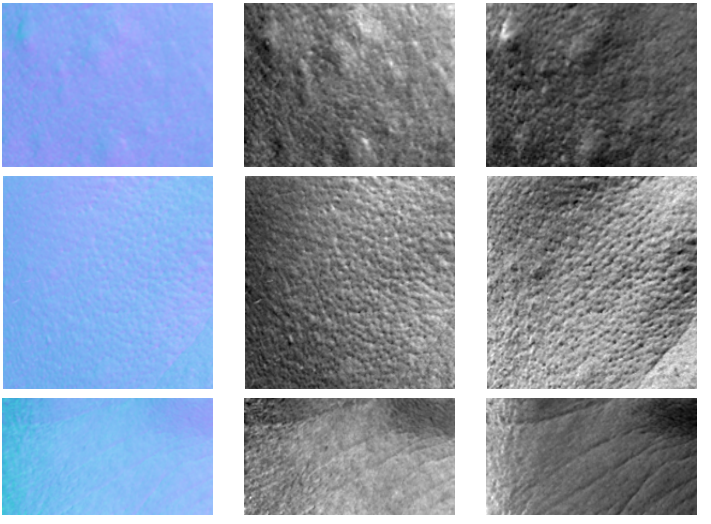


Figure 6: Comparison of APDI and our modified version (1st column: specular normal-map; 2nd column: APDI; 3rd column: our modified APDI)

3 Multi-Resolution Azimuthal Projection Distance Image

We present in this section a multi-scale scheme of APDIs for analysing 3D texture from dense surface orientations. For a given normal map, a multi-scale APDI pyramid is built by computing the normal map's APDIs at different resolutions. Let us firstly describe our geodesic-based normal map scaling algorithm.

3.1 Scaling Normal Maps

As stated earlier, normals do not lie in a linear space and therefore applying linear operations such as convolution directly on a normal map is not ideal. So we have to find a theoretically consistent way of rescaling normal maps. There are a number of methods for 3D mesh re-sampling [5, 10] which could be used to reduce or expand a mesh before recomputing the normals, but this is limited to cases where the geometry is known. Today photometric stereo techniques allow recovering surface orientations without depth, so methods for re-sampling normal maps independently to the geometry are needed. Here we propose a geodesic-based normal map re-sampling approach.

We use Riemannian differential geometry elements to introduce a new metric (geodesic distance) which will allow us to perform linear operations on the normals. We assume the normals to be on a Riemannian manifold and compute all linear operations on a tangent plane that we chose to be constant for all the normals. Let Exp_z and Log_z be the Riemannian Exponential and Logarithm operations [3] with z as projection axis, the linear combination of N normals $(n_i)_{1\leq i\leq N}$ with coefficients $(\alpha_i)_{1\leq i\leq N}$ can be computed as:

$$f(n_i,\alpha_i) = Exp_\mu\left(\sum_{i=1}^{N} \alpha_i \times Log_\mu(n_i)\right) \tag{7}$$

By definition from the Exponential mapping, the result will always be a unit vector. Our scaling algorithm is based on that equation. As we are only interested in down-sampling in this paper, we present in 1 an overview of the down-sampling algorithm. The full implementation includes border checking and index checking which we have omitted here for brevity.

Algorithm 1: Normal map down-sampling algorithm

Data: N: Normal map, S: Scale factor, $[u,v]$: Window size
Result: N': Down-sampled Normal map
$n_w \leftarrow \lfloor width(N)/S \rfloor$
$n_h \leftarrow \lfloor height(N)/S \rfloor$
for $i \leftarrow 1$ **to** n_w **do**
  for $j \leftarrow 1$ **to** n_h **do**
    $Tmp \leftarrow 0$
    for $w \leftarrow i - \lfloor \frac{u}{2} \rfloor$, $k \leftarrow 1$ **to** u **do**
      for $y \leftarrow j - \lfloor \frac{v}{2} \rfloor$, $l \leftarrow 1$ **to** v **do**
        $Tmp \leftarrow Tmp + Log_z(N(w,y))$
        $y \leftarrow y+1$
      end
      $w \leftarrow w+1$
    end
    $N'(i,j) \leftarrow Exp_z(\frac{Tmp}{u \times v})$
  end
end

We tested the proposed method by comparing a normal map with the result of down-sampling and up-sampling it back (we used the normal map on Figure 2). The geodesic method achieves 0.027 mean angular error, while using a classical sampling method on each channel and renormalizing back the result gives a mean angular error of 0.183.

3.2 Multi-Resolution APDI

With a consistent scaling algorithm, constructing the multi-resolution APDI pyramid becomes straightforward. Given a normal map, the pyramid is built by down-sampling the normal map to different levels. At each level, the APDI is computed from the corresponding down-sampled normal. The high levels contains higher frequency details adequate for texture analysis. The more we go down to the lower levels, the more we lose high frequency details, but the low frequency changes related to the overall shape are highlighted. An interesting way of capturing fine detail would be extracting the loss or the difference between levels, but this is out of the scope of this paper.

Figure 7 gives a 4-level APDI pyramid of different skin patches. At each level we show the (down-sampled) normal map used to compute the APDI.

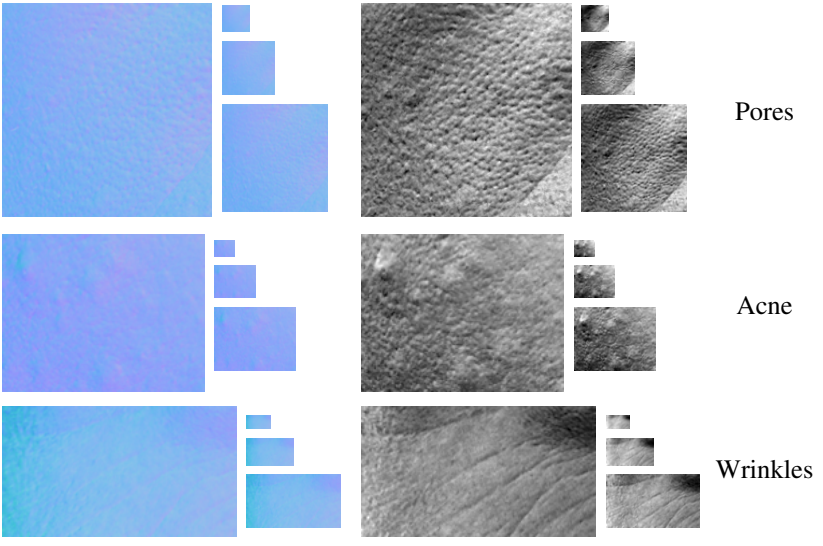


Figure 7: Multi-scale APDIs for different skin patches

4 Experimental Results

We used the proposed Multi-scale APDI and Local Binary Pattern to classify wrinkles and pores on high resolution facial normal maps. The experiment is preliminary as the dataset used (the ICT-3DRFE [13]) does not contain that much skin condition variation. We are in the process of collecting our own skin condition dataset using the same class of 3D capture device used on the ICT-3DRFE dataset collection (a Lightstage [7]) in order consider more skin conditions such as acne, moles etc.

Even though the ICT-3DRFE does not contain that much skin condition and ageing variation, we have managed to manually extract and rate on a scale of 4 levels (1 meaning "total absence" and 4 meaning "extremely visible") 31 patches visually judged wrinkly, 21 patches showing large pores and 30 smooth patches (for negative samples). The rating is done from photo-realistic rendering of the faces Figure 8 where the rater can select and give score to skin patches.

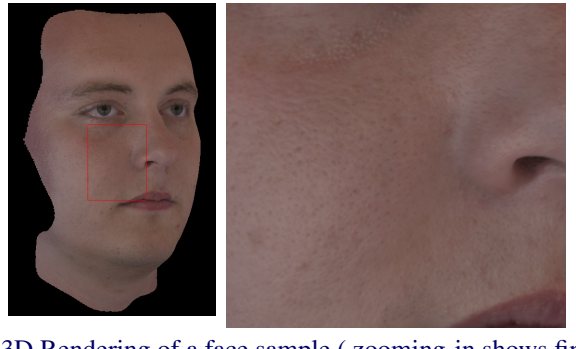


Figure 8: Our 3D Rendering of a face sample (zooming-in shows fine skin detail)

Each patch is divided in to a number of fixed-size blocks (with 30% overlap). This gives us a dataset with 384 rated skin samples. We then compute multi-scale APDI pyramid on each block. Local Binary Pattern features are finally extracted and concatenated from each level of the APDI pyramid. Concatenating the features from the different levels produces a relatively big feature vector depending on the number of levels and the LBP configuration (e.g. a 8-pixel neighbourhood and 1-pixel radius gives on a 4-level M-APDI a feature vector 1024 long). To reduce the feature dimensionality we rank the features with the method described in [4] which evaluates each feature's discriminative power using an SVM classifier. After the feature ranking, we select the 256 best features and use a Multilayer Perceptron classifier for learning and classification. Tables 1 and 2 summarize the classification result for different M-APDI pyramid height using respectively a 66-34% split and a 10-fold cross validation.

Skin Condition	APDI	2-level M-APDI	3-level M-APDI	4-level M-APDI
Wrinkle	0.786	0.826	0.930	0.965
Pore	0.858	0.943	0.971	0.988

Table 1: Results showing the accuracy of assigning the correct label :1(not visible), 2(slightly visible), 3(visible) or 4(extremely visible) wrinkles/pores to the patches. 66-34% SPLIT

Skin Condition	APDI	2-level M-APDI	3-level M-APDI	4-level M-APDI
Wrinkle	0.802	0.911	0.987	0.991
Pore	0.860	0.945	0.992	1.000

Table 2: Results showing the accuracy of assigning the correct label :1(not visible), 2(slightly visible), 3(visible) or 4(extremely visible) wrinkles/pores to the patches. 10-fold CROSS VALIDATION

5 Conclusion

We proposed a Multi-scale Azimuthal Projection Distance Image for characterising 3D surface texture from facial skin normal maps and applied it to automated skin condition assessment. We modified the original formulation of APDI in order to take account of the azimuthal orientation and introduced a method for resampling normal maps. Even though the dataset we have used in this preliminary experiment does not contain much skin condition variation, the results are very promising. We are working on collecting our own dataset and look forward to testing the proposed approach on a wider range of skin conditions.

References

[1] Chen H. and Bhanu B. 3D free-form object recognition in range images using local surface patches. *Pattern Recogn. Lett.*, 28(10):1252–1262, 2007.

[2] Ding Y., Smith L., Smith M., Sun J. and Warr R. Obtaining malignant melanoma indicators through statistical analysis of 3D skin surface disruptions. *Skin Research and Technology*, 15:262–270, 2009.

[3] Fletcher T., Lu C., Pizer S.M. and Joshi S. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions Medical Imaging*, 23:995–1005, 2004.

[4] Guyon I., Weston J., Barnhill S. and Vapnik V. Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning*, 46(1-3):389–422, 2002.

[5] Hoppe H. New quadric metric for simplifiying meshes with appearance attributes. In *Proceedings of the conference on Visualization '99: celebrating ten years*, VIS '99, pages 59–66, Los Alamitos, CA, USA, 1999. IEEE Computer Society Press. ISBN 0-7803-5897-X. URL http://dl.acm.org/citation.cfm?id=319351.319357.

[6] Johnson A. *Spin-Images: A Representation for 3-D Surface Matching*. PhD thesis, The Robotics Institute, Carnegie Mellon University, 1997.

[7] Ma W.C., Hawkins T., Peers P., Chabert C.F., Weiss M. and Debevec P. Rapid Acquisition of Specular and Diffuse Normal Maps from Polarized Spherical Gradient Illumination. In *Eurographics Symposium on Rendering*, 2007.

[8] Nehab D., Rusinkiewicz S., Davis J. and Ramamoorthi R. Efficiently Combining Positions and Normals for Precise 3D Geometry. *ACM Transactions on Graphics (Proc. of ACM SIGGRAPH 2005)*, 24(3), August 2005.

[9] Peyré G. and Mallat S.P. Surface compression with geometric bandelets. In *ACM SIGGRAPH*, volume 24, pages 601–608, 2005.

[10] Pietroni N., Tarini M. and Cignoni P. Almost isometric mesh parameterization through abstract domains. *IEEE Transaction on Visualization and Computer Graphics*, 16(4): 621–635, August 2010. URL http://vcg.isti.cnr.it/Publications/2010/PTC10.

[11] Sandbach G., Zafeiriou S. and Pantic M. Binary Pattern Analysis for 3D Facial Action Unit Detection. In *Proceedings of the British Machine Vision Conference*, 2012.

[12] Smith M., Anwar S. and Smith L. 3D Texture Analysis using Co-occurrence Matrix Feature for Classification. In *Fourth York Doctoral Symposium on Computer Science*, 2011.

[13] Stratou G., Ghosh A., Debevec P. and Morency L. Effect of illumination on automatic expression recognition: A novel 3D relightable facial database . *IEEE International Conference onAutomatic Face & Gesture Recognition and Workshops*, pages 611–618, 2011.

[14] Xavier P. Statistical Computing on Manifolds: From Riemannian Geometry to Computational Anatomy. In *in Emerging Trends in Visual Computing*.

[15] Zhong Y. Intrinsic shape signatures: A shape descriptor for 3D object recognition. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 689–696, 2009.