

Reconstructing 3D Facial Shape Using Spherical Harmonics

Chengming Zou^{1,2}, Guanghui Zhao¹, and Edwin R. Hancock²

¹ School of Computer Science, Wuhan University of Technology,
Wuhan, Hubei, 430070, China
{zoucm,zhaogh}@hotmail.com

² Department of Computer Science, The University of York, York, YO10 5DD, UK
{zoucm,erh}@cs.york.ac.uk

Abstract. It is now well established that 3D facial shape can be effectively and conveniently reconstructed using spherical harmonics. This paper extends the state-of-the-art by showing how to recover a 3D facial shape reconstruction using a spherical parameterization and minimizing a Harmonic Energy based on spherical medians. The solution is formulated as a linear system and we propose an iterative residual fitting algorithm (LSQR-IRF) to solve it. We demonstrate the effectiveness of the method on range-data.

Keywords: Spherical Harmonics, Reconstruction, 3D Facial Shape.

1 Introduction

Recently, it has been shown that statistical models based on the distribution of surface normals can offer a powerful means of representing an recognising facial shape. The reasons for this are two-fold. First, the needle map (or Gauss map) offers a representation that is rich in terms of differential geometry, and hence can potentially be used to capture subtle variations in facial shape. Secondly, surface reflectance is determined by the relative orientation of the surface and the light source. Hence, given a description of skin reflectance then a surface normal model can be fitted to image brightness data to recover 3D shape from a 2D facial image.

There are a number of approaches to capturing the statistics of surface normal direction. For instance, Smith and Hancock [1] project the surface normals into a tangent space to construct a statistical model using principal geodesic analysis. This work has recently been extended to gender recognition [2], but has proved too cumbersome for expression recognition. Bronstein, Bronstein and Kimmel [3] develop a spherical embedding, that allows faces to be represented in a manner that is invariant to expression. Parameterising the distribution of surface normals, Kazhdan et al. [4] use the fact that the spherical harmonics of a given frequency form a subspace which is a rotationally invariant and which can be applied to the extended gaussian image (EGI). to create a rotationally invariant shape descriptor.

In this paper we present a novel approach for 3D facial shape reconstruction using spherical harmonics. Our main contribution is to develop a robust spherical parameterization for mapping of facial shape meshes onto a unit sphere. The spherical harmonic representation decomposes polar functions into the frequency domain. Thus, we can reconstruct the 3D facial shape with a linear combination of spherical harmonics.

2 Spherical Harmonics

The spherical harmonic H_n^m of degree n and order m is defined as[5]

$$H_n^m(\theta, \phi) = \begin{cases} c_n^m P_n^{|m|}(\cos \theta) \sin(|m|\phi), & -n \leq m \leq -1 \\ \frac{c_n^m}{\sqrt{2}} P_n^m(\cos \theta), & m = 0 \\ c_n^m P_n^m(\cos \theta) \cos(m\phi), & 1 \leq m \leq n \end{cases} \quad (1)$$

where $c_n^m = \sqrt{\frac{2n+1}{2n} \frac{(n-|m|)!}{(n+|m|)!}}$ and the associated Legendre polynomials P_n^m are the solutions of the differential equation $P_n^m(x) = \frac{(-1)^n}{2^n n!} (1-x^2)^{\frac{m}{2}} \frac{d^{n+m}}{dx^{n+m}} (x^2-1)^n$.

Any spherical continuous function $f(\theta, \phi)$ can be represented by a linear combination of spherical harmonics $H_n^m(\theta, \phi)$ as follows:

$$f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n a_n^m H_n^m(\theta, \phi) \quad (2)$$

where the coefficients a_n^m are uniquely determined by[5]

$$a_n^m = \int_0^\pi \int_0^{2\pi} H_n^m(\theta, \phi)^* f(\theta, \phi) \sin \theta d\phi d\theta \quad (3)$$

where $H_n^m(\theta, \phi)^*$ is the complex conjugate of $H_n^m(\theta, \phi)$.

The spherical harmonic expansion described above is essentially the Fourier transform for functions defined on the sphere, and it transforms spherical scalar signals into the frequency domain. Spherical harmonics are orthonormal, complete, and can be used to construct a coarse-to-fine hierarchy. This makes them a natural choice as basis functions to reconstruct complex object surfaces including faces [6]. We denote a 3D facial surface mesh M as $s(\theta, \phi)$. In order to describe $s(\theta, \phi)$, we just need to expand these spherical functions using spherical harmonics[7]:

$$s(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_n^m H_n^m(\theta, \phi) \quad (4)$$

where $A_n^m = (a_{xn}^m, a_{yn}^m, a_{zn}^m)^T$ is vector encoding the Cartesian coordinates $a_{xn}^m, a_{yn}^m, a_{zn}^m$ of the vertices of M .

Spherical harmonics reconstruct a surface in a coarse-to-fine hierarchy. With increasing degree n the spherical harmonics seem to capture increasing detail. If we truncate the the maximal degree at N_{max} , the surface is approximated by

$$s(\theta, \phi) \approx \sum_{n=0}^{N_{max}} \sum_{m=-n}^n A_n^m H_n^m(\theta, \phi) \quad (5)$$

3 Spherical Parameterization

The first step in constructing spherical harmonics is to calculate the spherical signal as a function of θ and ϕ , and this process is referred to as spherical parameterization.

3.1 The Geodesic Distance

A unit vector $m \in R^3$ may be considered as a point lying on a spherical manifold $n \in S^2$, where S^2 is the unit sphere, and the two representations are related by $m = \Phi(n)$, where $\Phi : S^2 \rightarrow R^3$ is an embedding. If $v \in T_n S^2$ is a vector on the tangent plane to the manifold at $n \in S^2$ and $v \neq 0$, then the exponential map, denoted Exp_n , of v is the point on S^2 along the geodesic in the direction of v at distance $\|v\|$ from n . The inverse of the exponential map is the log map, denoted by Log_n . The geodesic distance between two points $n_1 \in S^2$ and $n_2 \in S^2$ can be expressed in terms of the log map, and is denoted by [1,8]

$$d(n_1, n_2) = \|Log_{n_1}(n_2)\| \quad (6)$$

Since a unit sphere is a Riemannian manifold and great circles are geodesics, arc length is the Riemannian distance between a pair of points and hence [1],

$$d(n_1, n_2) = \arccos(\Phi(n_1) \cdot \Phi(n_2)) \quad (7)$$

3.2 Spherical Medians

If we consider the distribution of unit vectors as a distribution of points on a spherical manifold $n_1, \dots, n_k \in S^2$, it is clear that the mean direction is dependent on the embedding Φ and is the extrinsic mean of a distribution of spherical data [1]:

$$\mu_\Phi = \arg \min_{n \in S^2} \sum_{i=1}^k \|\Phi(n) - \Phi(n_i)\|^2 \quad (8)$$

In other words, the extrinsic mean is the Euclidean average (or center of mass) of the distribution of points in R^3 , projected back onto the closest point on the sphere. A more natural definition of the average of a distribution of points on the unit sphere uses arc length as the choice of geodesic distance measure. Using (7), we can define the intrinsic mean [1]:

$$\mu = \arg \min_{n \in S^2} \sum_{i=1}^k d(n, n_i) \quad (9)$$

For spherical data, this is known as the spherical median.

3.3 Harmonic Energy

Suppose h is a homeomorphism between R^3 and S^2 , we define the harmonic energy as [9]:

$$E(h) = \sum_{[u,v] \in K} k_{u,v} \| h(u) - h(v) \|^2 \tag{10}$$

where K are the edges of M , and $k_{u,v}$ are the spring constants computed as in [10]. Given 3D facial shape data in the form of unit vectors, we use h to map these vectors onto a unit sphere. In order to avoid folding, h should be harmonic. If h is harmonic, then $E(h)$ acquire a minimal value. In other words, in order to diffuse h to harmonic, we should minimize $E(h)$.

For $k_{u,v}$ are constants, so in order to simplify calculation, we let $k_{u,v} \equiv 1$. We know that $h(u) \in S^2$, so $\| h(u) - h(v) \|$ is the geodesic distance the point $h(u)$ to the point $h(v)$. Thus minimizing $E(h)$ can be denoted as follow:

$$\arg \min \sum_{[u,v] \in K} d(n_{h(u)}, n_{h(v)})^2 \tag{11}$$

The edges $[u, v] \in K$ of the mesh form a 1-ring for the neighbors of u , denoted $Star(u)$. From (9), the intrinsic mean μ is such that the total distance between μ and $Star(u)$ will be minimum. For every u , we may calculate the intrinsic mean of $Star(u)$, and if we move u close to μ , then $\sum_{v \in Star(u)} d(n_{h(u)}, n_{h(v)})^2$ will be reduced. Thus if every u is the intrinsic mean of $Star(u)$, Eq. (11) will be minimum. Thus h is harmonic and can be used for spherical parameterization.

3.4 Spherical Parameterization Algorithm

In summary, we give our algorithm for spherical parameterization as follow:

Algorithm 1. Spherical Parameterization

Input: mesh M , step length δt

Output: spherical signal θ and ϕ

Step 1) Compute the spherical barycentric map b from M to S^2 , $h \leftarrow b$.

Step 2) Compute the harmonic energy $E(h)$, if $\delta E < \varepsilon$ then goto 7).

Step 3) Compute intrinsic mean μ of every point of h , and set $\Delta h = \delta t \cdot \mu$.

Step 4) Update h using $h \leftarrow h - \Delta h$.

Step 5) Shift the center of mass of $h(M)$ to origin and normalize $h(v), v \in K$ to the unit vector.

Step 6) Repeat 2) through 5).

Step 7) Compute Euler angles θ and ϕ by h .

Step 8) return θ and ϕ .

In order to ensure the convergence of the algorithm, Step 5) is added so that the solution is unique. It is executed to force the centre-of-mass of the surface to remain at the origin during entire update process.

4 Computing Spherical Harmonics

4.1 Calculating Associated Legendre Polynomial

In SPHARM, the efficient computation of the Associated Legendre Polynomial (ALP) is an important issue. One popular method is the Recursive Algorithm (RA) which easy to implement, but it is not efficient. To overcome this limitation we use a non recursive algorithm. We note the following properties of $P_n^m(x)$ From the definition of ALP, we know that $P_n^m(x) = 0$ when $n < |m|$ and from Ref.[11]:

$$P_n^m(x) = ((2n - 1)xP_{n-1}^m(x) + (n + m - 1)P_{n-2}^m(x))/(n - m) \tag{12}$$

$$P_{n+1}^{n+1}(x) = -(2n + 1)(1 - x^2)^{1/2}P_n^n(x) \tag{13}$$

$$P_n^{n-1}(x) = (2n - 1)xP_{n-1}^{n-1}(x) \tag{14}$$

Based on these properties our algorithm for calculating ALP as follow (where $|x| < 1$ is:

Algorithm 2. Calculate Associated Legendre Polynomial

Input: x, n , where $|x| < 1, n > 0$

Output: $P_0^0, P_1^0, P_1^1, \dots, P_n^0, P_n^1, \dots, P_n^n$

Step 1) Set $P_0^0(x) = 1.0$.

Step 2) Use formula (13) to calculate $P_1^1(x), P_2^2(x), \dots, P_n^n(x)$.

Step 3) Use formula (14) to calculate $P_1^0(x), P_2^1(x), \dots, P_n^{n-1}(x)$.

Step 4) Use the result of 2) and 3) and formula (12) to calculate $P_n^m, \forall |m| \leq n-2$.

This is a non-recursive algorithm and although it uses more memory than the recursive algorithm, for a given x and degree n , it can calculate all the $P_n^m(x)$ in one pass with greatly improved the efficiency.

4.2 LSQR[12] Iterative Residual Fitting

We focus our discussion on expanding a spherical scalar signal $f(\theta, \phi)$. For the SPHARM case, we can apply the same method three times and expand $x(\theta, \phi)$, $y(\theta, \phi)$ and $z(\theta, \phi)$ separately. Given a function $f(\theta, \phi)$ and a user-specified maximum degree L_{max} , we can reconstruct the original function as follows:

$$\hat{f}(\theta, \phi) = \sum_{l=0}^{L_{max}} \sum_{m=-l}^l \hat{a}_l^m Y_l^m(\theta, \phi) \approx f(\theta, \phi) \tag{15}$$

The higher the degrees the more accurate the reconstruction $\hat{f}(\theta, \phi)$. Unfortunately, with increasing L_{max} , the above linear system becomes intractable. To overcome this problem we exploit the properties of the spherical harmonic transform. First, we make use of the coarse-to-fine hierarchy provided by SPHARM's. If we use a just a few low degree SPHARM's to expand a spherical function

$f(\theta, \phi)$, we obtain a low-pass filtered reconstruction. If we use a higher degree representation, then the reconstruction is more detailed. Taking advantage of this coarse-to-fine hierarchy, we start from a low degree reconstruction and then iteratively add higher degree SPHARM's to increase the reconstructed detail. Since, SPHARM's form an orthonormal basis the reconstruction residual from a low frequency representation (i.e., $f(\theta, \phi)$ -its reconstruction) encode the unreconstructed high frequency detail. To add in more detail to our model, we can simply use a few higher degree harmonics to fit the residual.

To develop this idea, we first introduce some notation:

$$A_n = \begin{pmatrix} Y_n^{-n}(\theta_1, \phi_1) & Y_n^{-n+1}(\theta_1, \phi_1) & \cdots & Y_n^n(\theta_1, \phi_1) \\ Y_n^{-n}(\theta_2, \phi_2) & Y_n^{-n+1}(\theta_2, \phi_2) & \cdots & Y_n^n(\theta_2, \phi_2) \\ \vdots & \vdots & \vdots & \vdots \\ Y_n^{-n}(\theta_w, \phi_w) & Y_n^{-n+1}(\theta_w, \phi_w) & \cdots & Y_n^n(\theta_w, \phi_w) \end{pmatrix} \tag{16}$$

$$b_n = (\hat{a}_n^{-n}, \hat{a}_n^{-n+1}, \dots, \hat{a}_n^n)^T, M = (b_0^T, b_1^T, \dots, b_{L_{max}}^T)^T \tag{17}$$

$$f = (f(\theta_1, \phi_1), f(\theta_2, \phi_2), \dots, f(\theta_w, \phi_w))^T \tag{18}$$

where A_n is a matrix of degree n spherical harmonic values for w sampling points, b_n contains the estimated degree n coefficients, f is a vector of spherical function values at w sampling points, and M is a spherical harmonic model containing coefficients of up to degree L_{max} . Note that the model M minimizes

$$\sum_{i=1}^w (f(\theta_i, \phi_i) - \sum_{n=0}^{L_{max}} \sum_{m=-n}^n \hat{a}_n^m Y_n^m(\theta_i, \phi_i))^2 \tag{19}$$

We can now formally present our algorithm as follows:

Algorithm 3. LQSR Iterative Residual Fitting(LQSR-IRF)

Input: A spherical function f with w sampled values $\{\theta_i, \phi_i, f(\theta_i, \phi_i) | 1 \leq i \leq w\}$, the maximal SH expansion degree L_{max}

Output: Reconstruct model M

Step 1) $r = (f(\theta_1, \phi_1), f(\theta_2, \phi_2), \dots, f(\theta_w, \phi_w))^T$

Step 2) $M = (0, 0, \dots, 0)^T \{Initialization, |M| = (1 + L_{max})^2\}$

Step 3) Solving $A_0 x_0 = r$ by LSQR get x'_0 , set $b_0 = x'_0, r = r - A_0 x'_0$

Step 4) Set $d = 1$

Step 5) while $d \leq L_{max}$ do

Step 6) Solving $A_d x_d = r$ by LSQR get x'_d , set $b_d = x'_d, r = r - A_d x'_d$

Step 7) Set $d = d + 1$

Step 8) Set $M = M + (b_0^T b_1^T \dots b_{L_{max}}^T)^T$

Step 9) return model M

5 Experiments

We have performed extensive experiments on several 3D facial range data-sets. Our algorithm is implemented in VC2005. The experiments are performed on a computer with a 1GHZ AMD Athlon Processor and 1GB of RAM, running WinXP. We have performed three experiments. The first is the reconstruction of face01 with a neutral facial expression which degree=58. The second is reconstruction face01 with six facial expressions, which degree=58. The final experiment is a reconstruction of face02 with neutral facial expression using in varying degree, and specifically degree=5,18,28,32,38,48,58,68.

5.1 Results

Each of the experimental runs completes in 5 minutes. The reconstructions obtained are shown in Figure 1 to Figure 3.



Fig. 1. Reconstruction face01 with neutral facial expression, left is source



Fig. 2. Reconstruction face01 with six facial expressions



Fig. 3. Reconstruction face02 with neutral facial expression in several degrees

5.2 Modeling Accuracy

Let $f(\theta, \phi)$ be raw 3D facial range image and $s(\theta, \phi)$ be its SPHARM reconstruction. In order to measure the SPHARM modelling accuracy, we calculate the mean squared distance (msd) and the maximal squared distance (xsd) between the n mesh vertices on the original surface and their reconstructions as follows.

$$msd = \frac{1}{w} \sum_{i=1}^w \| f(\theta_i, \phi_i) - s(\theta_i, \phi_i) \|^2 \quad (20)$$

$$xsd = \max\{\| f(\theta_i, \phi_i) - s(\theta_i, \phi_i) \|^2 \mid 1 \leq i \leq w\} \quad (21)$$

Table 1. The *msd* and *xsd*

Experiment	<i>msd</i>	<i>xsd</i>	<i>Degree</i>	<i>Expression</i>
1	0.000523	0.003588	58	neutral
2	0.000523	0.003588	58	neutral
2	0.000767	0.006114	58	smile
2	0.001041	0.013198	58	laugh
2	0.000301	0.002516	58	sad
2	0.000279	0.067474	58	surprised
2	0.000178	0.004020	58	pleased
3	0.001323	0.020878	5	neutral
3	0.000837	0.012049	18	neutral
3	0.000781	0.007028	28	neutral
3	0.000772	0.006274	32	neutral
3	0.000764	0.005928	38	neutral
3	0.000759	0.006263	48	neutral
3	0.000756	0.006256	58	neutral
3	0.000755	0.006058	68	neutral

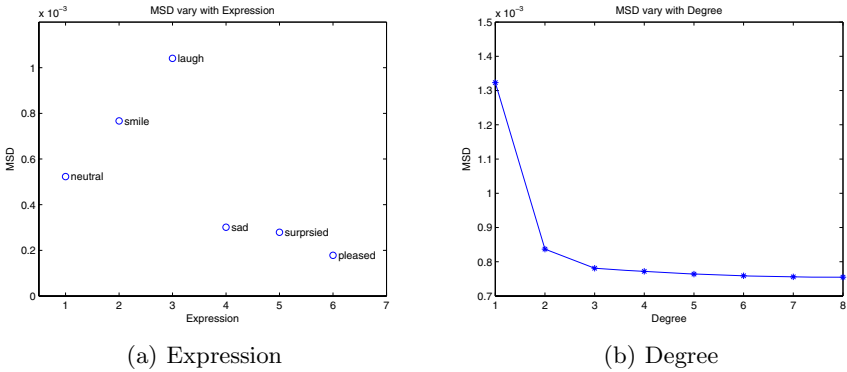


Fig. 4. MSD varying with expression and maximum degree of spherical harmonics

The values of *msd* and *xsd* found in our experiments are given in Table 1 and Figure 4. From Table 1 and Figure 4, it is clear that the reconstruction is accurate and we can reconstruct different facial expressions. As expected where the degree increases, the error decreases. Finally Figure 4 explores the structure of the error. The left-hand plot shows the embedding obtained, when multidimensional scaling is applied to the Euclidean distance between the vectors of fitted SPHARM co-efficients for the different expressions. The embedding reveals that the co-efficient distance increases with increasing departure from the neutral expression. The right-hand plot of Figure 4 shows the MSD as a function of the degree of the fitted SPARM model. As expected, the error decreases with increasing model order.

6 Conclusions

We have presented a novel framework for 3D facial shape reconstruction based on spherical harmonics. We used the intrinsic mean and harmonic map to implement a spherical parameterization. We provided an efficient non-recursive algorithm for calculating the associated Legendre polynomials. We also gave an iterative residual fitting method (LQSR-IRF) for reconstructing facial shape using spherical harmonics. Experimental results have shown that our method for 3D facial shape reconstruction is feasible and effective.

References

1. Smith, W.A.P., Hancock, E.R.: Facial Shape-from-shading and Recognition Using Principal Geodesic Analysis and Robust Statistics. *Int. Journal Computer Vision* 76, 71–91 (2008)
2. Wu, J., Smith, W.A.P., Hancock, E.R.: Weighted principal geodesic analysis for facial gender classification. In: Rueda, L., Mery, D., Kittler, J. (eds.) *CIARP 2007*. LNCS, vol. 4756, pp. 331–339. Springer, Heidelberg (2007)
3. Bronstein, A., Bronstein, M., Kimmel, R.: Expression invariant face recognition via spherical embedding. In: *Image Processing, ICIP*, vol. 3, pp. 756–759 (2005)
4. Kazhdan, M., Funkhouser, T., Rusinkiewicz, S.: Rotation Invariant Spherical harmonic representation of 3D shape descriptors. *Proceedings of the 2003 Eurographics* 43, 126–164 (2003)
5. Chung, M.K., Dalton, K.M., Shen, L., et al.: Weighted Fourier Series Representation and Its Application to Quantifying the Amount of Gray Matter. *IEEE Transactions on Medical Imaging* 26(4), 566–581 (2007)
6. Ballard, D.H., Brown, C.M.: *Computer Vision*. Prentice-Hall, Englewood Cliffs (1982)
7. Shen, L., Chung, M.K.: Large-Scale Modeling of Parametric Surfaces using Spherical harmonics. In: *Third International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)* (2006)
8. Fletcher, P.T., Lu, C., Pizer, S.M., et al.: Principal Geodesic Analysis for the Study of Nonlinear Statistics of Shape. *IEEE Transactions on Medical Imaging* 23(8), 995–1005 (2004)
9. Gu, X., Yau, S.-T.: Computing Conformal Structures of Surfaces. *Communications in Information and Systems* 2(2), 121–146 (2002)
10. Eck, M., DeRose, T., Duchamp, T., et al.: Multiresolution Analysis of Arbitrary Meshes. In: *Proceedings of SIGGRAPH 1995*, pp. 173–182 (1995)
11. Ye, Q., Shen, Y.: *Practical Handbook of Mathematics*, 2nd edn. Science Press, China (2006)
12. Paige, C.C., Saunders, M.A.: LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. *ACM Transactions on Mathematical Software* 8(1), 43–71 (1982)