SHAPE-FROM-SHADING: COMBINING PROBABILISTIC & STATISTICAL MODELS

ICVSS 2010
Registration and Video Analysis

Ahmad T.[1], Wilson R.C., Department of Computer Science, University of York, UK

{touqeer,wilson}@cs.york.ac.uk

ABSTRACT

In this work, we combine a probabilistic model of surface normals from shape-from-shading[1] with a statistical model of 3D shape[2]. We sample the Fisher Bingham FB8 distribution of surface normals from probabilistic SfS model using Gibbs sampling. We get surface normals from the statistical shape model. We fit individual normal distributions to each of these normals and combine them to give a product normal distribution i.e. a better model for SfS.

INTRODUCTION

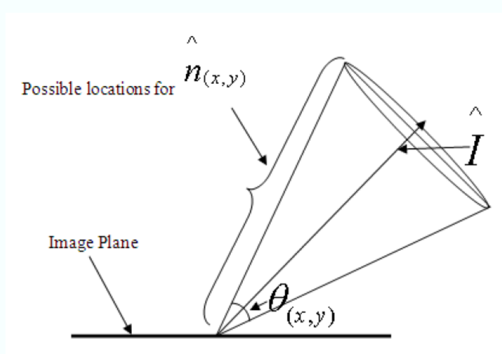
'To get 3D shape from a single 2D gray scale image' i.e. Shape-from-Shading as it is commonly known in Computer Vision community; has been a major research topic from past four decades. In last forty years a number of approaches [4] have been proposed to improve the solution. Recently an approach [1] based on directional statistics and belief propagation has been presented that have shown better results as compared to other currently available algorithms [2-4]. However Smith's work [2] has shown great improvement over geometrical SfS algorithm [3] when a 3D statistical model of human face shapes was used. Inspired by the improvements shown by incorporation of 3D statistical model we are going to combine the probabilistic model of SFS [1] with this 3D shape model based on the Yale-B data base [6].

PROBABILISTIC SFS

Lambertian shading equation applies;

$$I_{x,y} = a\hat{\mathbf{l}} \cdot \hat{\mathbf{n}}_{x,y} \qquad \frac{I_{x,y}}{a} = \cos\theta_{x,y}$$

The normal direction has two degrees of freedom i.e. elevation angle and the azimuth angle. The elevation angle is given by the cone constraint. This leaves one degree of freedom per pixel. The algorithm takes a conventional approach and uses smoothness as further source of information.



There were two reasons to choose the FB8 distribution – firstly the multiplication of two FB8 distributions gives another FB8 distribution and secondly the sub-model of Bingham-Mardia distribution is used to represent the cone constraint. Each pixel has a prior expressing the cone constraint and the compatibility between the adjacent pixels provide the smoothness term. A graphical model i.e. a pair-wise Markov random field on a grid is constructed in which each node is a random variable that represents the unknown surface orientation of a pixel. Sum-product belief propagation is used to determine the marginal distribution for each node i.e. a FB8 distribution.

STATISTICAL SHAPE MODEL

If $v \in T_nS^2$ is a vector on the tangent plane to S^2 at $n \in S^2$ and $v \neq 0$ the exponential map denoted by Exp_n of v is the point on S^2 along the geodesic in the direction of v at distance ||v|| from n. The point on S^2 is denoted by $Exp_n(v)$. The inverse of the exponential map is the log map denoted by Log_n. The log and exponential maps are given by,

$$Exp_n(v) = \left(v_1 . \frac{\sin\|v\|}{\|v\|}, v_2 . \frac{\sin\|v\|}{\|v\|}, \cos\|v\|\right)$$

$$Log_n(x) = \left(x_1 . \frac{\theta}{\sin\theta}, x_2 . \frac{\theta}{\sin\theta}\right) \quad \text{where; } \theta = \arccos(x_3)$$

Intrinsic mean of set of points on spherical manifold is a point that minimizes the Riemannian distance to each point in the given set. This minimization problem is solved iteratively by gradient descent method of Pennec [7].

$$\mu_{(j+1)} = \mathrm{Exp}_{\mu_{(j)}}\left(\frac{1}{K}\sum_{i=1}^{K}\mathrm{Log}_{\mu_{(j)}}(n_i)\right)$$

SAMPLING of FB8

We use the slice sampling algorithm proposed by A.Kume & S.G.Walker [5] to sample from the Fisher Bingham distribution. The Fisher Bingham distribution is the multivariate normal distribution having a constraint to lie on the surface of a sphere. If **x**={x0, x1, x2} is distributed according to FB distribution then |**x**|=1. Due to this constrain $\mathbf{x}^2=\{x_0^2, x_1^2, x_2^2\}$ lies on the simplex; so **x** is transformed to the (**w**, **s**) and a joint density of (**w**, **s**) is found using Gibbs method by introducing latent variables (*u*, *v*, *w*). Here, $s_i=x_i^2$ and $w_i=x_i/|x_i|$. When transformed rom **x** to ($\mathbf{w}_0$, $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{s}_1$, $\mathbf{s}_2$) we get the joint density of interest.

$$f(\boldsymbol{\omega}, \mathbf{s}) \propto \exp\left\{\sum_{i=1}^{p}(a_i s_i + b_i\omega_i\sqrt{s_i})\right\} \times \exp\left\{b_0\omega_0\sqrt{1-s}\right\} \times \prod_{i=1}^{p} s_i^{-1/2}(1-s)^{-1/2}\mathbf{1}(s \le 1).$$

The three latent variables (*u*, *v*, *w*) are Introduced to give the joint density as,

$$f(\boldsymbol{\omega}, \mathbf{s}, u, v, w) \propto \mathbf{1}\{u < \exp(L)\}\mathbf{1}\{v < \exp(b_0\omega_0\sqrt{1-s})\} \times \mathbf{1}\{w < (1-s)^{-1/2}\}\prod_{i=1}^{p} s_i^{-1/2}\mathbf{1}(s \le 1),$$

The three latent variables are uniform and easy to sample from; the conditional densities of **w** and **s** are then found accordingly.

$$\mathrm{P}(\omega_i = +1|\cdots) = \frac{\exp(b_i\sqrt{s_i})}{\exp(-b_i\sqrt{s_i}) + \exp(b_i\sqrt{s_i})}$$

$$f(s_1|\cdots) \propto \mathbf{1}\{A_u \cap A_v \cap A_w \cap A\}\, s_1^{-1/2},$$

The sets *A*, A_u, A_w and A_v are formed by inverting the inequalities involving latent variables.

EXPERIMENT DETAILS

We calculate the intrinsic mean for each pixel location using 250 pre-aligned face shapes using the method described above. We use five iterations to get the intrinsic mean. The covariance matrix Σ_1 for each pixel location is also calculated from a matrix of 250 x 3 surface normals for that particular pixel location. We define a normal distribution for each pixel location(x, y) by $\Omega_1[P_1\mu_1, P_1]$ where $P_1=\Sigma_1^{-1}$ is inverse of the covariance matrix μ_1 is the intrinsic mean. From SfS we get Fisher Bingham distributions for each pixel, we apply the Gibbs sampling algorithm described above and get samples from each distribution. Through five iterations and 250 samples per iteration we get the intrinsic mean and covariance matrix for each pixel location (x, y) and fit a normal distribution $\Omega_2[P_2\mu_2, P_2]$. The two distributions are then multiplied to give the most probable surface normal μ^* at pixel location (x, y).

$$\Omega_T[P^*\mu^*, P^*] = \Omega_1[P_1\mu_1, P_1]\,\Omega_2[P_2\mu_2, P_2]$$

$$\Omega_T[P^*\mu^*, P^*] = \Omega[P_1\mu_1 + P_2\mu_2, P_1 + P_2]$$

$$\mu^* = (P^*)^{-1}(P^*\mu^*)$$

RESULTS

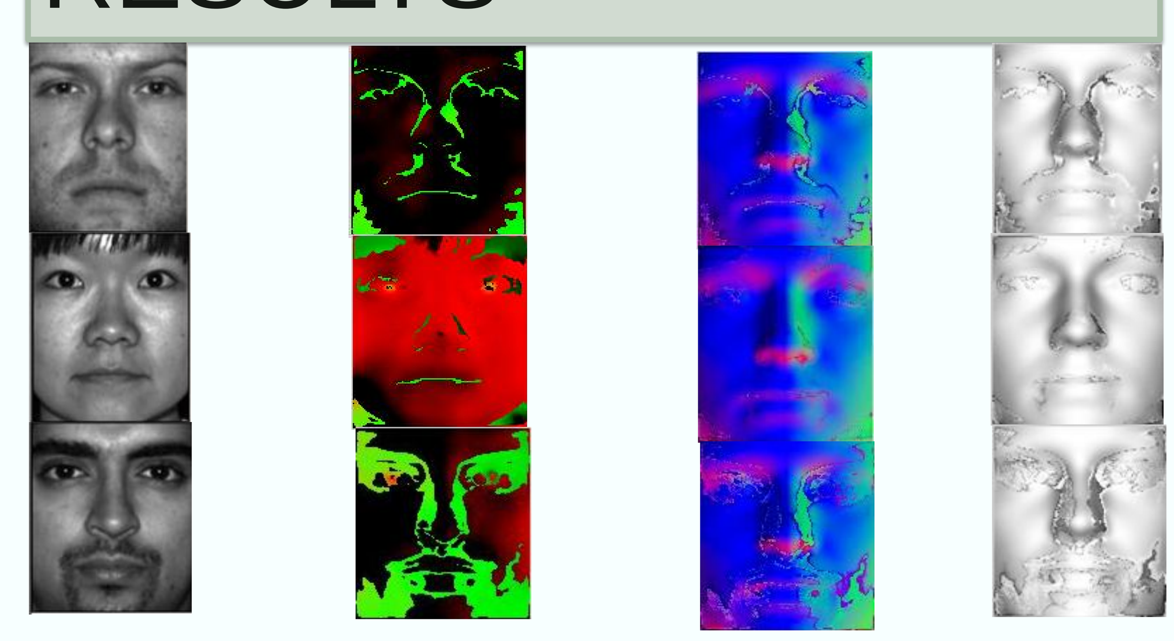


Image | SfS Needle Map | Combined Needle Map | Re-Illuminated

Preliminary results have shown improvement over the SfS model even when the surface normals are combined through normal distributions. In coming months we will try to fit some more suitable distributions and iterate the method to get better results.

ACKNOWLEDGEMENT

Touqeer Ahmad is thankful to European Research and Educational Collaboration with Asia - EURECA scholarship programme for fully funding his M.Sc. in Computer Science (by research) at University of York.

REFERENCES

[1] Tom S.F. Haines, *Integrating Shape-from-Shading & Stereopsis*, Ph.D. Dissertation, University of York, York, UK, Sep. 2009.
[2] William A.P. Smith, *Statistical Methods for Facial Shape-from-Shading & Recognition*, Ph.D. Dissertation, University of York, York, UK, Feb. 2007.
[3] P. L. Worthington and E. R. Hancok: New Constraints on Data-Closeness and Needle Map Consistency for Shape-from-Shading. *IEEE Trans. Pattern Analysis and Machine Intell.*, 21(12):1250-1267, 1999.
[4] R. Zhang, P.S. Tsai, J. E. Cryer, and M. Shah: Shape-from-shading: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(8):690–706, 1999.
[5] A. Kume and S.G. Walker: On the Fisher-Bingham distribution. *Statistic and Computing*, 19(2):167-172, 2009.
[6] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman: From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose, *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(6): 643-660, 2001.
[7] X. Pennec: Probabilities and statistics on Riemannian manifolds: basic tools for geometric measurements. In Proc. *IEEE Workshop on Nonlinear Signal and Image Processing,* 1999.