Bias-Variance Tradeoff for Adaptive Surface Meshes

Richard C. Wilson and Edwin R. Hancock

Department of Computer Science
University of York
York, Y01 5DD, UK
email: wilson,erh@minster.york.ac.uk

Abstract. This paper presents a novel statistical methodology for exerting control over adaptive surface meshes. The work builds on a recently reported adaptive mesh which uses split and merge operations to control the distribution of planar or quadric surface patches. Hitherto, we have used the target variance of the patch fit residuals as a control criterion. The novelty of the work reported in this paper is to focus on the variance-bias tradeoff that exists between the size of the fitted patches and their associated parameter variances. In particular, we provide an analysis which shows that there is an optimal patch area which minimises the variance in the fitted patch parameters. This area offers the best compromise between the noise-variance, which decreases with increasing area, and the model-bias, which increases in a polynomial manner with area. The computed optimal areas of the local surface patches are used to exert control over the facets of the adaptive mesh. We use a series of split and merge operations to distribute the faces of the mesh so that each resembles as closely as possible its optimal area. In this way the mesh automatically selects its own model-order by adjusting the number of control-points or nodes. We provide experiments on both real and synthetic data. This experimentation demonstrates that our mesh is capable of efficiently representing high curvature surface detail.

1 Introduction

Adaptive meshes [3, 18, 19] have proved popular in both the segmentation [25, 12, 4, 13, 16, 8] and efficient representation [6] of volumetric surface data . The literature is rich with examples. For instance De Floriani et al [6, 7] have developed a multi-scale mesh which has been exploited not only for surface representation, but also for stereoscopic reconstruction. Several authors have reported variable topology meshes. Bulpitt and Efford [1] have a mesh that adapts itself so as to minimise curvature and goodness of fit criteria. The "slime" surface of Stoddart *et al* [20] uses region merge operations of refine a B-spline mesh surface.

These surfaces are effectively driven by geometric criteria [18, 19]. In a recent series of papers we have developed a surface mesh which is statistically motivated [27, 26]. Each node in our mesh represents a local quadric patch that is fitted

to a support neighbourhood on the surface. Specifically, we have shown how a series of node split and merge operations can be used to both refine and decimate the mesh so as to deliver a surface of predefined target variance. These operations not only control the surface topology, they also iteratively modify the support neighbourhoods for the quadric patch representation. Analysis of the mesh reveals that the equilibrium distribution of mesh nodes is such that the density is proportional to the underlying curvature of the surface. The surface has been demonstrated to produce useful segmentations that can be used for subsequent differential analysis [26].

The aim in this paper is to focus more closely on the statistical criterion that underpins the control of the mesh split and merge operations. In particular we consider the variance-bias tradeoff [10] which underpins the choice of the support neighbourhood for the estimation of surface parameters. Simple split-and-merge operations based on a target 'goodness of fit' can result in biased or noisy patch estimates. This has undesirable effects on the recovered differential structure of the surface [17, 22, 23]. It is for this reason that we present a detailed analysis of parameter variance. The main conclusion of this analysis is that the variance has a two-component structure. The first component results from the effects of noise and decreases with increasing area of estimation. The second term results from the model-bias and increases with the area of estimation. As a result of the interplay between these two terms, there is an optimal choice of the area of estimation that results in a joint minimisation of both the noise variance of the estimated parameters and the model bias.

The optimal local area of estimation is used to exert control over the split [19] and merge [18] operations that underpin our adaptive mesh. By driving the adaptation of the mesh from the optimal local patch area we provide a natural means of controlling the model-order for our surface representation. Using these split and merge operations, a mesh is generated which has faces of area equal to the optimal area of estimation. If surface patches are placed at each of these faces, the subsequent piecewise representation is optimal in the sense that the error to the underlying surface parameters is minimal. The parameters of the patch are sufficient to represent the surface to within the accuracy limits imposed by the noise.

2 Approximating the surface from noisy data-points

Following Besl and Jain [2] our aim is to fit increasingly complex variable order surface models to potentially noisy data-points. Viewed from the perspective of local surface geometry this can be viewed as sequentially estimating derivatives of increasing order through the fitting of an appropriate surface patch. When couched in this intrinsically hierarchical way, each derivative relies on the estimation of the preceding and lower order derivatives. As a concrete example, in order to estimate curvature through a second-order quadric patch, we must first fit zero and first order models to determine the surface height and surface normal direction. Subject to the limitations imposed by the level of image noise, this

process can obviously be extended to any model-order to estimate the desired derivative.

In practice however, there is a problem of variance-bias tradeoff that hinders the parameter estimation process. By increasing the size of the sample or surface area used to estimate the model parameters, the effects of noise variance may be minimised. In other words, the temptation is to increase the size of the local surface patches so as to increase the accuracy of the estimated derivatives. Unfortunately, as the surface facet is increased in area problems of model bias emerge. In a nutshell, the problem is that the model order is insufficient to represent genuine structure in the data. The basic issue addressed in this paper is how to resolve this dilemma for the important and generic problem of adaptive mesh control. It must be stressed that variance-bias issues are ones of pivotal philosophical and practical importance in data fitting [10].

We commence our discussion with a set of 3-dimensional data-points $P = \{\mathbf{p}_i | \forall i\}$ derived from range data. In realistic tasks, these points are invariable uncertain in the sense that they deviate from the true surface due to some noise process. In the following, we denote the function of the underlying surface as $f(x, y)$ and the equation of points on this underlying surface is therefore $z = f(x, y)$. The data-point $\mathbf{p}_i = (x_i, y_i, z_i)$ is related to the true surface by $z_i = f(x_i, y_i) + n_i$, where n_i is the additive noise process. Now consider the Taylor expansion of the true surface function.

$$f(x_i, y_i) = f(x_o, y_o) + \left\{\frac{\partial f}{\partial x}\right\}_o (x_i - x_o) + \left\{\frac{\partial f}{\partial y}\right\}_o (y_i - y_o) + \left\{\frac{\partial^2 f}{\partial x^2}\right\}_o (x_i - x_o)^2 + \left\{\frac{\partial^2 f}{\partial x \partial y}\right\}_o (x_i - x_o)(y_i - y_o) + \left\{\frac{\partial^2 f}{\partial y^2}\right\}_o (y_i - y_o)^2 \tag{1}$$

If we wish to estimate, for example, first order derivatives (corresponding to the surface normal), we must first estimate the height $z_o = f(x_o, y_o)$ and remove it's contribution to the Taylor expansion by moving the origin of the coordinate system to $(x_o, y_o, f[x_o, y_o])$. The derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial x}$ can the be estimated by fitting the tangent plane $z = ax + by$. Similarly, the surface curvature can be determined by transforming co-ordinates to remove zero and first order contributions. The necessary second-order derivatives are estimated by fitting the quadric patch $f(x, y) = \alpha x^2 + \beta xy + \gamma y^2$ to the transformed height data.

In the remainder of this section, we provide an analysis of the errors in the fitted local surface models for each derivative in turn, commencing with the estimation of surface location. In each case we provide an analysis of variance for the fitted surface parameters. This commences from the known variances of the surface fit-residuals. These residuals are propagated through into the estimation of surface parameter variances. This analysis reveals the area dependence of the parameter variances. It is this analysis which allows us to estimate the optimal patch-area which results in the best variance-bias tradeoff.

2.1 Estimating the average surface height

We first consider the estimation of the height of the local origin of co-ordinates on the surface. We denote this point by the vector of co-ordinates $\boldsymbol{o} = (x_o, y_o, z_o)^T$. This location of the origin can be estimated by the mean height of the data-points. If S denotes the index-set of the sample of available points and n_i represents the additive noise present in the height measurement z_i, then

$$\boldsymbol{o} = \frac{1}{|S|}\sum_{i\in S}\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \frac{1}{|S|}\sum_{i\in S}\begin{pmatrix} x_i \\ y_i \\ f(x_i, y_i) + n_i \end{pmatrix} \tag{2}$$

If the sampled points are uniformly distributed over $x-y$ footprint of the surface patch, then the x and y co-ordinates of the origin are located at the centre-of-mass of the support neighbourhood. The height distribution, on the other hand, is governed both by the sampling noise and the bias introduced by the underlying surface shape in the sampling window. The two processes have very different origins. The noise process is stochastic and requires an explicit statistical model. The bias is a measure of the inappropriateness of the surface shape model adopted in the local sampling window.

The contributions from the various sources can be evaluated by again using the Taylor expansion. Our estimate of the local location of the surface is given by the average height of the sample data-points thus;

$$z_{est} = \frac{1}{|S|}\sum_{i\in S}[n_i + z_o + \Delta x_i\frac{\partial f}{\partial x} + \Delta y_i\frac{\partial f}{\partial y} + \Delta x_i^2\frac{\partial^2 f}{\partial x^2} + \Delta x_i\Delta y_i\frac{\partial^2 f}{\partial x\partial y} + \Delta y_i^2\frac{\partial^2 f}{\partial y^2} \tag{3}$$

where $\Delta x_i = x_i - x_o$ and $\Delta y_i = y_i - y_o$. Because we have chosen a symmetrical sampling window, the odd spatial moments $\sum_{i\in S}\Delta x_i$, $\sum_{i\in S}\Delta y_i$ and $\sum_{i\in S}\Delta x_i\Delta y_i$ are zero. As a result, the estimated height-intercept is given by

$$z_{est} = z_o + \frac{1}{|S|}\sum_{i\in S} n_i + \frac{1}{|S|}\frac{\partial^2 f}{\partial x^2}\sum_{i\in S}\Delta x_i^2 + \frac{1}{|S|}\frac{\partial^2 f}{\partial y^2}\sum_{i\in S}\Delta y_i^2 + \ldots \tag{4}$$

In other words, the estimated height-intercept is deviates from the average z-value by an amount that is determined by the second-order derivatives of the surface. The variance of the fit-residuals, i.e. $\sigma_{est}^2 = \frac{1}{|S|}\sum_{i\in S}(z_i - z_o)^2$ therefore has a two-component structure. The first of these results from averaging the raw image noise over the $|S|$ samples in the local surface patch. When the noise n_i is assumed to follow a Gaussian distribution with variance σ^2 and zero mean, then the average noise variance is equal to $\frac{1}{|S|}\sigma^2$. The second contribution to the variance of the fitted height originates from the derivative bias terms, of which the most significant terms are the second order derivatives of the surface. As a result, the total variance is given by

$$\sigma_{est}^2 = \frac{\sigma^2}{|S|} + \left\{\frac{\partial^2 f}{\partial x^2}\right\}_o^2\left(\sum_{i\in S}\Delta x_i^2\right)^2 + \left\{\frac{\partial^2 f}{\partial y^2}\right\}_o^2\left(\sum_{i\in S}\Delta y_i^2\right)^2$$

$$+ \left\{ \frac{\partial^2 f}{\partial x^2} \right\}_o \left\{ \frac{\partial^2 f}{\partial y^2} \right\}_o \sum_{i \in S} \Delta x_i^2 \sum_{i \in S} \Delta y_i^2 \tag{5}$$

If the data-points are uniformly distributed over the x-y footprint of the surface patch, then the expectation values of the second-order moments $\sum_{i \in S} \Delta x_i^2$ and $\sum_{i \in S} \Delta y_i^2$ can be calculated from the geometric moments of the support neighbourhood. Of course, the exact values of these moments depend on both the size and shape of the support neighbourhood. For simplicity, we will evaluate the moments for a circular region around the origin. In order to make the role of the area A of the support neighbourhood explicit, we replace the number of points in the sample-set by the expression $|S| = \rho A$ where ρ is the surface-density of data-points. As a result, the expectation value for $\sum_{i \in S} \Delta x_i^2$ is given, in the circular case, by

$$E(\sum_{i \in S} \Delta x_i^2) = \rho \int_A x^2 dA = \frac{\rho A^2}{8\pi} \tag{6}$$

Substituting for the expectation values of the second-order surface-moments, the final expression for the total variance is given by

$$\begin{aligned} \sigma(A)^2_{est} &= \frac{\sigma^2}{\rho A} + \frac{A^2}{64\pi^2} \left[\left\{ \frac{\partial^2 f}{\partial x^2} \right\}_o^2 + \left\{ \frac{\partial^2 f}{\partial x^2} \right\}_o \left\{ \frac{\partial^2 f}{\partial y^2} \right\}_o + \left\{ \frac{\partial^2 f}{\partial y^2} \right\}_o^2 \right] \\ &= \frac{\sigma^2}{\rho A} + k_0 A^2 \end{aligned}$$

The two component area dependence of the total variance is now made explicit. The first term represents the propagation of raw noise variance. As the area of the support neighbourhood used in the estimation of the origin increases, then so the effect of noise-variance on the fitted parameters decreases. The second term, on the other hand, represents the model bias in the extracted parameters. In the case of estimating the origin, the bias depends on the second derivatives, or curvature, of the local surface. The bias term increases with increasing surface area. It is clear that there is critical value of the area which results in minimum total variance. We locate the minimum area by fitting an empirical model to the measured height variance observed for various support neighbourhood areas. This fitting process returns estimates of the two model parameters σ^2 and k_0. When these have been extracted from the the variance-area data, the optimal area is given by

$$A_{min} = \left(\frac{\sigma^2}{2\rho k_0} \right)^{\frac{1}{3}} \tag{7}$$

In the Sections 2.2 and 2.3, we extend our analysis to the estimation of local surface orientation and curvature. In both cases, there is a similar variance-bias structure to the total variance.

2.2 Estimation of the surface normal

By translating to the local system of co-ordinates centred on the origin o, we we can again perform the Taylor expansion for the local surface patch. In the translated co-ordinate system,

$$z' = \left\{\frac{\partial f}{\partial x}\right\}_o x' + \left\{\frac{\partial f}{\partial y}\right\}_o y' + \left\{\frac{\partial^2 f}{\partial x^2}\right\}_o x'^2 + \left\{\frac{\partial^2 f}{\partial x \partial y}\right\}_o x'y' + \left\{\frac{\partial^2 f}{\partial y^2}\right\}_o y'^2 \quad (8)$$

where $x' = x - x_o$ and $y' = y - y_o$ are the translated co-ordinates. In the transformed co-ordinate system the height intercept of the local surface patch is zero. Hence, the tangent-plane may be estimated directly from the data-points in this new coordinate system.

Both the parameter estimation process and the propagation of variance is more complicated than in the case of the origin. Parameter estimation is realised by the least-squares fit of a tangent plane through the origin $z' = ax' + by'$. Again we choose a set of sample data-points S. We denote the parameters of the tangent plane by $\mathbf{P} = (a, b)^T$. The positions of the sample points are represented by the design matrix

$$\mathbf{X}_p = \begin{pmatrix} x'_1 & y'_1 \\ x'_2 & y'_2 \\ \vdots & \vdots \end{pmatrix}$$

while the corresponding height data is represented by the column-vector $\mathbf{Z}_p = (z'_1, z'_2, ...)^T$. The least-squares fit for the parameters is given by $\hat{\mathbf{P}} = \mathbf{L}_p \mathbf{Z}_p$ where $\mathbf{L}_p = (\mathbf{X}_p^T \mathbf{X}_p)^{-1} \mathbf{X}_p^T$ is the pseudo-inverse of the design matrix.

When the parameter-vector $\mathbf{P}$ is estimated in this way, then its covariance structure can be found by propagating the variance in the transformed height data $\mathbf{Z}_p$. If $\mathbf{\Sigma}_{Z_p}$ is the covariance matrix for the transformed height data, then the the covariance matrix for the plane parameters, i.e. $E[(\mathbf{P} - \hat{\mathbf{P}})(\mathbf{P} - \hat{\mathbf{P}})^T]$, is given by

$$\mathbf{\Sigma}_p = \mathbf{L}_p \mathbf{\Sigma}_{Z_p} \mathbf{L}_p^T \quad (9)$$

As in the case of the origin, the total covariance matrix has a two-component structure which reflects the two sources of error in the estimation of the surface normals. The first component is due to the propagation of noise in the surface-data-point positions, while the second component is a bias term that results from the higher order terms in the Taylor expansion. We make this two-component structure more explicit by writing

$$\mathbf{\Sigma}_p = \mathbf{L}_p \mathbf{\Sigma}_N \mathbf{L}_p^T + \mathbf{L}_p \mathbf{\Sigma}_B \mathbf{L}_p^T \quad (10)$$

The noise component of the parameter covariance matrix is modelled under the assumption that transformed height data is subject to independent identically distributed Gaussian noise of zero mean and variance σ^2. Under this assumption the noise variance of the least-squares parameter estimates is given by

$$\mathbf{L}_p \mathbf{\Sigma}_N \mathbf{L}_p^T = \sigma^2 \begin{pmatrix} \sum_{i \in S} x'^2_i & \sum_{i \in S} x'_i y'_i \\ \sum_{i \in S} x'_i y'_i & \sum_{i \in S} y'^2_i \end{pmatrix}^{-1} \quad (11)$$

In other words, the noise-component to the total covariance matrix depends on the second-order moments of the points in the surface patch. As before we assume a circular support neighbourhood. In this case, the expectation values of the odd co-ordinate moments are zero. The expectation values of the even moments can be computed along the same lines as outlined in the previous subsection. As a result the noise contribution has a diagonal covariance matrix Specifically,

$$\mathbf{L}_p\mathbf{\Sigma}_N\mathbf{L}_p^T = \frac{12\sigma^2}{\rho A}I \tag{12}$$

where I is the 2x2 identity matrix.

The bias contribution is more complex and depends, as before, on the second-order, and higher, derivatives of the local surface. We model the bias term to second-order by computing the covariance matrix for the local deviations from the planar approximation. Accordingly, we write bias-component of the covariance matrix as

$$\mathbf{\Sigma_B} = \begin{pmatrix} \epsilon_1\epsilon_1 & \epsilon_1\epsilon_2 & \dots & \epsilon_1\epsilon_n \\ \epsilon_2\epsilon_1 & \ddots & & \\ \vdots & & & \end{pmatrix} \tag{13}$$

where $\epsilon_i = \frac{\partial^2 f}{\partial x^2}(x_i - x_o)^2 + \frac{\partial f}{\partial x}\frac{\partial f}{\partial y}(x_i - x_o)(y_i - y_o) + \frac{\partial^2 f}{\partial y^2}(y_i - y_o)^2 + \dots$ is the non-planar deviation of the point indexed i.

Details of the bias model are outside the scope of this paper. Suffice to say that, we can compute the expectation values for the elements of the non-planar bias covariance matrix in much the same way as for the case of estimating the patch height, neglecting higher order terms of the expansion. Under this condition, the bias can be represented as a second-order polynomial in the patch area A. If $\mathbf{K}_0$, $\mathbf{K}_1$ and $\mathbf{K}_2$ represent co-efficient matrices whose elements depend on the second order and higher derivatives of the surface function, then

$$\mathbf{L}_p\mathbf{\Sigma}_B\mathbf{L}_p^T = \mathbf{K}_0 + \mathbf{K}_1 A + \mathbf{K}_2 A^2 + \dots \tag{14}$$

Collecting together terms, we find that the total parameter covariance matrix can be expressed as

$$\mathbf{\Sigma}_p(A) = \frac{12\sigma^2}{\rho A}\mathbf{I} + \mathbf{K}_0 + \mathbf{K}_1 A + \mathbf{K}_2 A^2 + \dots \tag{15}$$

Again, the noise propagation term is inversely proportional to the area of the estimating patch. The bias terms, on the other hand, are polynomial in area. As a result the parameter covariance matrix can be minimised with respect to the patch area.

The problem of determining the optimal area of estimation for surface normals is more complicated than in the case of the average height. The main difficulty stems from the fact that we are dealing with a covariance matrix rather than a single scalar quantity. However, since the noise component of $\mathbf{\Sigma_p}$ is diagonal, we confine our attention to minimising the trace of the covariance matrix.

To first order in area, the trace is given by

$$Tr[\mathbf{\Sigma_p}] = \sigma_a^2 + \sigma_b^2 = 2\left[(\frac{12\sigma^2}{\rho A} + k_0 + k_1 A)\right] \tag{16}$$

where σ_a^2 and σ_b^2 are the measured variances for the plane parameters a and b. Again, we can fit the predicted area dependance to the observed sum of variances $\sigma_a^2 + \sigma_b^2$ to estimate the semi-empirical parameters σ, k_0 and k_1. The minimum error surface patch area is given by

$$A_{min} = \left[\frac{\rho k_1}{12\sigma^2}\right]^{\frac{1}{2}} \tag{17}$$

2.3 Estimation of the surface curvature

The estimation of surface curvature proceeds in much the same way as for the surface normals. We begin by transforming the coordinate system in such a way as to remove both zero order and first order terms of the Taylor expansion. From a geometric perspective this is equivalent to translating the origin and rotating into the local tangent plane. If the local co-ordinate system is located at the point $\boldsymbol{o}$, then the z-axis of co-ordinates is directed along the surface normal. The x and y axes are orthogonal to one-another and are oriented arbitrarily in the local tangent plane of the surface. In the local coordinate system, the Taylor expansion is now given by

$$f(x'', y'') = \left\{\frac{\partial^2 f}{\partial x''^2}\right\}_{\boldsymbol{o}} x''^2 + \left\{\frac{\partial^2 f}{\partial x'' \partial y''}\right\}_{\boldsymbol{o}} x''y'' + \left\{\frac{\partial^2 f}{\partial y''^2}\right\}_{\boldsymbol{o}} y''^2 + O(x''^3) \tag{18}$$

It is now clear that the natural approximate representation of this surface in the local coordinate system is a quadric patch

$$f(x'', y'') = \alpha x''^2 + \beta x''y'' + \gamma y''^2 \tag{19}$$

In other words, the vector of parameters $\mathbf{Q} = (\alpha, \beta, \gamma)^T$ represents the estimate of the second order derivatives of the surface around the origin $\boldsymbol{o}$ of the local coordinate system. We obtain estimates of the parameters $\hat{\mathbf{Q}} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})^T$ using least-squares fitting over the raw data-points that associate with a support neighbourhood of area A on the surface. The solution vector is given by $\hat{\mathbf{Q}} = \mathbf{L}_q \mathbf{Z}_q$ where the design matrix of transformed sample-points is given by

$$\mathbf{X}_q = \begin{pmatrix} x''^2_1 & x''_1 y''_1 & y''^2_1 \\ x''^2_2 & x''_2 y''_2 & y''^2_2 \\ & \vdots & \end{pmatrix}$$

and the transformed height data is now represented by the column-vector $\mathbf{Z}_q = (z''_1, z''_2, ...)^T$. Again the pseudo-inverse of the design matrix $\mathbf{L}_q$ is given by $\mathbf{L}_q = (\mathbf{X_q^T X_q})^{-1} \mathbf{X_q^T}$.

The parameter covariance matrix again has a variance-bias structure. We make this explicit by writing

$$\mathbf{\Sigma}_Q = \mathbf{L}_q \mathbf{\Sigma}_N \mathbf{L}_q^T + \mathbf{L}_q \mathbf{\Sigma}_B \mathbf{L}_q^T \tag{20}$$

The covariance component originating from additive noise is related to the fourth-order moments of the x and y co-ordinates in the support neighbourhood. Specifically, we find

$$\mathbf{L}_q \mathbf{\Sigma}_N \mathbf{L}_q^T = \sigma^2 \begin{pmatrix} \sum_{i \in S} x_i''^{\,4} & \sum_{i \in S} x_i''^{\,3} y_i'' & \sum_{i \in S} x_i''^{\,2} y_i''^{\,2} \\ \sum_{i \in S} x_i''^{\,3} y_i'' & \sum_{i \in S} x_i''^{\,2} y_i''^{\,2} & \sum_{i \in S} x_i'' y_i''^{\,3} \\ \sum_{i \in S} x_i''^{\,2} y_i''^{\,2} & \sum_{i \in S} x_i'' y_i''^{\,3} & \sum_{i \in S} y_i''^{\,4} \end{pmatrix}^{-1} \tag{21}$$

The expectation values of the matrix elements can be estimated as before, and are given by

$$< \sum_{i \in S} x_i''^{\,4} >= \rho \int_A x''^{\,4} dA = \frac{\rho A^3}{8\pi^2} \tag{22}$$

and

$$< \sum_{i \in S} x_i''^{\,2} y_i''^{\,2} >= \rho \int_A x''^{\,2} y''^{\,2} dA = \frac{\rho A^3}{24\pi^2} \tag{23}$$

The expectation-values for the remaining fourth-order moments which involve odd-powers of x or y are zero. Hence we may write the noise component of the covariance matrix explicitly in terms of the area of the support neighbourhood in the following manner

$$\mathbf{L_q \Sigma}_N \mathbf{L}_q^T = \frac{3\pi^2\sigma^2}{\rho A^3} \begin{pmatrix} 3 & 0 & -1 \\ 0 & 8 & 0 \\ -1 & 0 & 3 \end{pmatrix} = \frac{3\pi^2\sigma^2}{\rho A^3} \mathbf{K_N} \tag{24}$$

Details of the analysis of the bias in the deviations from the local quadratic is more complicated and beyond the scope of this paper. Suffice to say that the bias component can be expanded in terms of a polynomial in A in the following manner

$$\mathbf{L}_q \mathbf{\Sigma}_B \mathbf{L}_q^T = \mathbf{K}_0 + A\mathbf{K}_1 + A^2\mathbf{K}_2 + \ldots \tag{25}$$

where $\mathbf{K}_0$ and $\mathbf{K}_1$ are matrices whose elements depend on the third order derivatives of the surface.

We can now write the covariance matrix of the quadric patch parameters in terms of the area of estimation (to first order terms) thus:

$$\mathbf{\Sigma}_Q = \frac{3\pi^2\sigma^2}{\rho A^3} \mathbf{K}_N + \mathbf{K}_0 + A\mathbf{K}_1 \tag{26}$$

Since the noise covariance matrix, i.e. $\mathbf{L}_q \mathbf{\Sigma}_N \mathbf{L}_q^T$ is no longer diagonal, we can no-longer strictly recover the optimal patch area by minimising the trace of $\mathbf{\Sigma}_Q$. However, from equation (24) it is clear that the off-diagonal elements are insignificant compared to the trace. Therefore, in the case of the quadric patch

parameters we approximate the optimal surface area by fitting a semi-empirical model to the measured trace of the quadric patch covariance matrix $\mathbf{\Sigma}_Q$.

$$Tr[\mathbf{\Sigma}_Q] = \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 = \left[\frac{42\pi^2\sigma^2}{\rho A^3} + k_0 + Ak_1\right] \tag{27}$$

where σ_α^2, σ_β^2 and σ_γ^2 are the measured variances for each of the quadric patch parameters in turn. As before, we calculate the values of the semi-empirical parameters σ^2, k_0 and k_1 by fitting the predicted area dependance to the measured variances. In this case the optimal surface area which minimises $Tr[\mathbf{\Sigma_q}]$ is given by

$$A_{min} = \left[\frac{\rho k_1}{14\pi^2\sigma^2}\right]^{\frac{1}{4}} \tag{28}$$

3 Controlling the Mesh

In the previous section, we provided an analysis of variance for the sequential extraction of a local surface origin, the local tangent plane, and, finally, the local patch curvature parameters. In each case we demonstrated that the fit covariance parameters could be minimised with respect to the area of the sample neighbourhood. In other words, there is an optimal choice of estimation area. This area can be viewed as providing the best tradeoff between model bias and underlying data variance. Our overall aim in this paper is to exploit this property to control the area of surface patches in an adaptive surface mesh. Since the area of the patches determines the number of nodes needed to represent the surface, the use of the optimal local patch area effectively corresponds to controlling the model-order of the surface representation. The interpretation of the mesh depends on the quantity which is being estimated over the surface. For example, in the case of surface normals, the centre of each face represents a point at which the surface normal is sampled. The surface is represented by the piecewise combination of the tangent planes associated with these sample normals. The mesh adapts itself to the data using a series of split and merge operations which are aimed at delivering a mesh which optimally represents the surface.

3.1 Optimal area estimation

Here we aim to use the minimum parameter-covariance area to control the split and merge operations. We directly estimate the optimal local patch-size and adjust the mesh topology accordingly. In an ideal world, the optimal area could be determined by varying the area of estimation and noting the value that minimises the parameter covariances. However there is an obstacle to the direct implementation of this process. In practice, the random nature of the noise component of the data-points results in multiple local minima.

The bias-variance relationships developed in the previous section allow us to overcome this difficulty. In particular, they suggest the overall model-dependance

between the variance in model parameters and the area of estimation. By fitting this semi-empirical model to the computed parameter variances we can smoothly estimate the position of the global minimum corresponding to the optimal area. The strategy that we adopt in determining the optimal local patch area is as follows. For each point on the surface we gradually increase the local patch area and compute the associated parameter variances. This gives a set of data points to which we can fit an appropriate empirical form of the bias-variance curve. The fitted parameters can be used to extract the value of the minimum local patch-area in a stable manner.

3.2 Mesh Adaptation

The optimal areas of estimation vary over the surface, and suggest that the level of representation, i.e. the model-order, required by the surface should also vary in line with the variance-bias criteria outlined in Section 2. To achieve this goal, we will adopt an adaptive mesh representation of the surface. In this representation, nodes of the mesh represent salient points on the surface; the distance between these points is such that there is the best trade-off between noise and bias in the positions of the points. The mesh points then represent the minimal accurate representation of the surface.

Our mesh is based on the Delaunay triangulation of a set of control points or nodes [24, 18, 19, 8, 9, 6]. In contrast to the bulk of the work reported in the literature which focus on the optimal positioning of the nodes [8, 24], in this paper it is the triangular faces of the mesh to which we turn our attention. The basic update process underpinning our surface involves adjusting the mesh-topology by splitting and merging surface-triangles. This process is realised by either inserting or deleting nodes from the mesh. The net effect of the two operations is to modify the node, edge and face sets of the mesh. The node insertion and deletion operations take place with the objective of delivering a set of faces whose areas are consistent with the optimal values dictated by the bias-variance criteria outlined in section 2. In this way the density of nodes is such as to strike a compromise between over-fitting the data and over-smoothing genuine surface detail. In other words, we seek the minimum model-order (i.e. the total number of nodes) such that each of the triangular faces is as close as possible to its optimal area.

Triangle merging is realised by deleting a node from the mesh as illustrated in the left-hand panel of Figure 1. The basic aim is to merge triangles if the aggregate area is more consistent with the optimal area than the original area. Suppose that the set of triangles M_j is to be merged to form a new triangle with area A_j. The average area of the configuration of triangles is

$$A_j^{merge} = \frac{1}{|M_j|} \sum_{i \in M_j} A_i \tag{29}$$

The triangles are merged if the fractional difference between the average area and the optimal area is greater than 10%. In other words, we instantiate the

merge if

$$\frac{A_j^{optimal} - A_j^{merge}}{A_j^{optimal}} > 0.1 \tag{30}$$

This tolerancing can be viewed as providing the adaptation of the mesh with a degree of hysteresis.

The geometry of the split operation is illustrated in the right-hand panel Figure 1. A new node is introduced at the centroid of the original triangle. The new node-set is re-triangulated to update the edge and face sets of the triangulation. The condition for initiating a split operation is that the current fractional difference between the triangle area and it optimal value is greater than 10%. The split condition can therefore be stated as

$$\frac{A_j - A_j^{optimal}}{A_j^{optimal}} > 0.1 \tag{31}$$

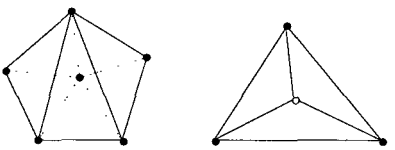


Fig. 1. Merging and splitting triangles

4 Experiments

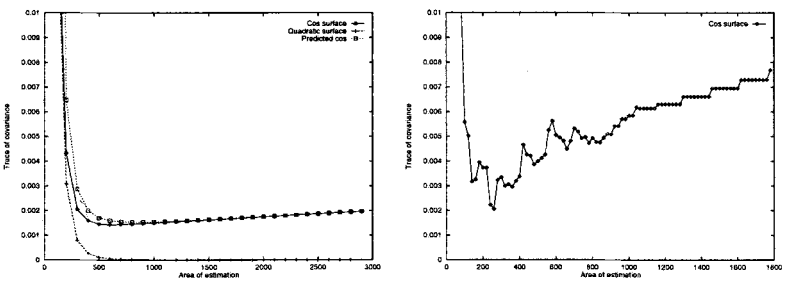
In this Section we provide some experimental evaluation of our new bias-variance controlled surface. There are several aspects to this study. We commence by considering synthetic data. The aim here is to illustrate that the two-component variance model described in Section 3 does provide a good description for the distribution of the parameter variances as a function of patch-area. The second aspect of our experimentation focuses on real world data-sets. Here we consider range-data from the Michigan State University data-base.

4.1 Synthetic Data

The main aim under this heading is to illustrate that the variance-bias description presented in Section 3 provides an accurate model of the distribution of the fitted parameter variances under controlled conditions. Moreover, we also aim to show that the simple parametric models developed in Section 4.1 are capable of fitting the observed distribution of variance as a function of the area used in the estimation process.

Figure 2a shows a series of plots of the parameter variance as a function of the area of estimation. Here we investigate the effect of quadric patch fitting on two different synthetic surfaces. The points marked with a cross show the computed variances when the surface is itself quadric. Here the parameter variances follow a distribution which monotonically decays with area. In other words, because the model is well matched to the data, the bias component of the parameter variance is zero and there is no local minimum area. The diamonds, on the other hand, show the variance for a surface of the form $f(z) = \cos(x)\cos(y)$. This surface is not well matched to the quadric patch approximation, and we may anticipate the model-bias term to emerge. This is evident from the fact that the variance shows a local minimum when the estimating area is approximately 500 points. In order to investigate the effectiveness of our analysis of variance in describing the observed distribution, the boxes show the result of plotting the prediction of the model outlined in Section 2.3. Although the agreement between the two curves is by no means perfect, the main features are captured by the model. Most importantly for the successful implementation of our adaptive mesh, there is good agreement in the location of the minimum variance area.

To illustrate the behavior of the parameter variance on realistic data, Figure 2b shows an example distribution for a range-image. Here there is genuine data-point noise and there is considerably more structure than in the case of the synthetic surface. However, the distribution maintains the same gross structure. There is a clear decaying component due to the surface noise together with an increasing component due to the bias term. The interplay between these two components results in a well defined global minimum. In other words, fitting an appropriately parameterised distribution should meet with success when attempting to estimate the minimum variance area.



(a) Synthetic surface. (b) Real-world data.

Fig. 2. Trace of the curvature covariance matrix for simulated surface.

4.2 Real World Data

In this Section we provide experiments on real world data-sets from the Michigan State University range-image archive. We commence by showing a sequence of results for a range image of a Renault part. Figure 3a shows the surface normals estimated using the local fitting process outlined in Section 3.1. The associated patches are shown in Figure 3b. In Figure 3c we shown the rendered surface patches. The main points to note from this sequence of images are as follows. Firstly, the extracted surface normals provide a faithful representation of the high curvature surfaces details. This is particularly evident around the sharp machined edges of the part. The second point to note is the distribution of surface triangles. In the flat portions of the image these are sparse. By contrast, in the highly curved regions of the surface, the density increases to account for the local curvature of the surface.

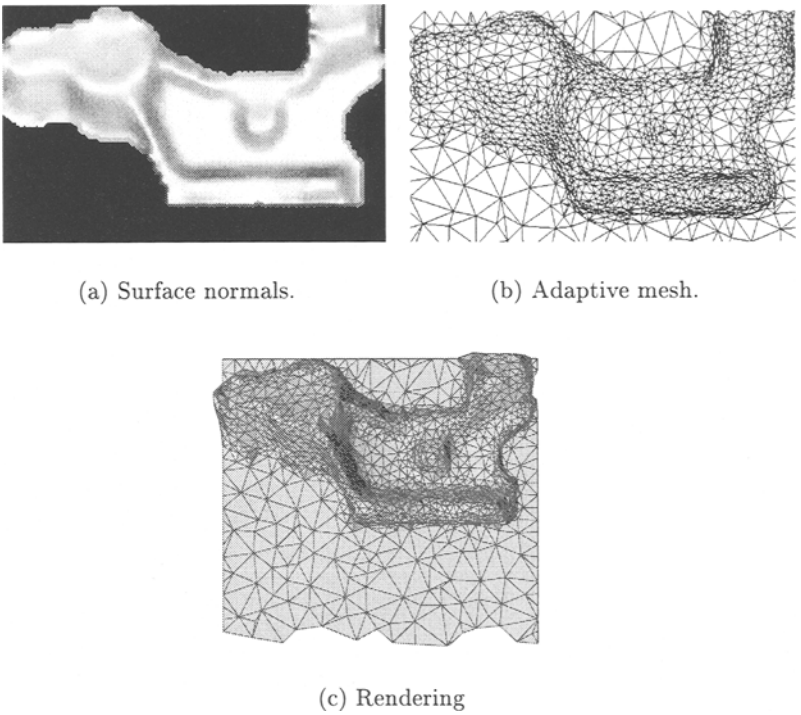
Figures 4a-4c show an analogous sequence of images for a range-image of a bust of the composer Mozart. This range-image contains considerably more fine detail than the Renault-part. The main point to note is the effectiveness of the mesh at representing areas of different curvature with mesh points of varying density.

The planar patch information is refined in Figure 4d where we show the mean curvature extracted from the subsequent quadric patch fit. The mean curvature is given $K = \alpha + \gamma$ where α and γ are the coefficients of the quadric patch as defined in equation 19. The estimated curvatures are signed. The maximum negative and positive values appear as extreme light and dark regions in the figure. The important feature to note from the figure is the fine detail in both the cravat and collar of the bust. There is also well defined detail around the concave features of the face (i.e. the eye sockets, the lips and the ridge of the nose).

5 Conclusions

The main contribution in this paper have been twofold. In the first instance, we have provided an analysis of variance for the various stages in establishing the derivatives of surface fits to range data. This analysis reveals that there is a two component structure to the variance of the fitted parameters. The first component of results from the propagation of noise variance and decreases with the increasing area of the patch. The second variance contribution originates from the model bias. This term is polynomial in the patch area. The main conclusion is that there is a local patch area which offers optimal tradeoff between noise-variance propagation and model bias in the minimisation of the fit residuals.

Our second contribution is to exploit this property to control the area of the facets of a triangulated surface mesh. By locally varying the patch area we minimise the parameter variance. This minimisation is realised by splitting and merging the triangular faces of the mesh until they each have a near optimal area.



(a) Surface normals. (b) Adaptive mesh.

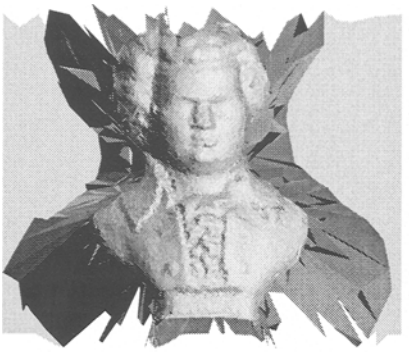
(c) Rendering

Fig. 3. Analysis of the Renault part.

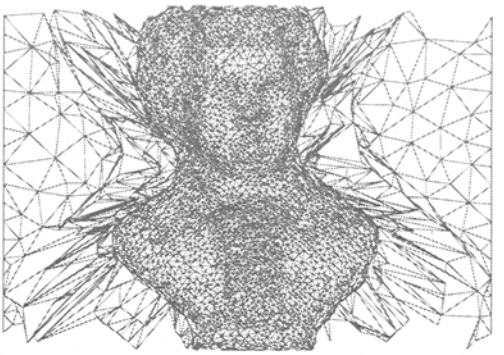
References

1. A.J. Bulpitt and N.D. Efford, "An efficient 3d deformable model with a self-optimising mesh", *Image and Vision Computing*, **14**, pp. 573–580, 1996.
2. P. J. Besl and R. C. Jain, "Segmentation through variable order surface fitting", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **10**:2 pp167-192 1988
3. X. Chen and F. Schmill. Surface modeling of range data by constrained triangulation", *Computer-Aided Design*, **26**, pp. 632—645 1994.
4. I. Cohen, L.D. Cohen and N. Ayache, "Using deformable surfaces to segment 3D images and infer differential structure", *CVGIP*, **56**, pp. 243–263, 1993.
5. I. Cohen and L.D. Cohen, "A hybrid hyper-quadric model for 2-d and 3-d data fitting", *Computer Vision and Image Understanding*, **63**, pp. 527–541, 1996.
6. L. De Floriani, " A pyramidal data structure for triangle-based surface description", *IEEE Computer Graphics and Applications*, **9**, pp. 67–78, 1987.
7. L. De Floriani, P. Marzano and E. Puppo, " Multiresolution models for topographic surface description.", *Visual Computer*, **12**:7, pp. 317–345, 1996.
8. H. Delingette, "Adaptive and deformable models based on simplex meshes", *IEEE Computer Society Workshop on Motion of Non-rigid and Articulated Objects*, pp. 152–157, 1994.

9. H. Delingette, M. Hebert and K. Ikeuchi, "Shape representation and image segmentation using deformable surfaces", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 467–472, 1991.
10. D. Geman, E. Bienenstock and R. Doursat, "Neural networks and the bias variance dilemma", *Neural Computation,*, **4**, pp.1-58, 1992.
11. J. O. Lachaud and A. Montanvert, "Volumic segmentation using hierarchical representation and triangulates surface", *Computer Vision, ECCV'96, Edited by B. Buxton and R. Cipolla, Lecture Notes in Computer Science, Volume 1064*, pp. 137–146, 1996.
12. R. Lengagne, P. Fua and O. Monga, "Using crest-lines to guide surface reconstruction from stereo", *Proceedings of th 13th International Conference on Pattern Recognition*, Volume A, pp. 9–13, 1996.
13. D. McInerney and D. Terzopoulos, "A finite element model for 3D shape reconstruction and non-rigid motion tracking", *Fourth International Conference on Computer Vision*, pp. 518–532, 1993.
14. D. McInerney and D. Terzopoulos, "A dynamic finite-element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image-analysis", *Computerised Medical Imaging and Graphics*, **19**, pp. 69–83, 1995.
15. M. Moshfeghi, S. Ranganath and K. Nawyn, "Three dimensional elastic matching of volumes", *IEEE Transactions on Image Processing*, **3**, pp. 128–138,
16. W. Neuenschwander, P. Fua, G. Szekely and O. Kubler, "Deformable Velcro Surfaces", *Fifth International Conference on Computer Vision*, pp. 828–833, 1995.
17. P.T. Sander and S.W. Zucker, "Inferring surface structure and differential structure from 3D images", IEEE PAMI, **12**, pp 833-854, 1990.
18. F. J. M. Schmitt, B. A. Barsky and Wen-Hui Du, "An adaptive subdivision method for surface fitting from sampled data", *SIGGRAPH '86*, **20**, pp. 176–188, 1986
19. W.J.Schroeder, J.A. Zarge and W.E. Lorenson, "Decimation of triangle meshes", *Computer Graphics*, **26** pp. 163–169, 1992.
20. A.J. Stoddart, A. Hilton and J. Illingworth, "SLIME: A new deformable surface", *Proceedings British Machine Vision Conference*, pp. 285–294, 1994.
21. A. J. Stoddart, J. Illingworth and T. Windeatt, "Optimal Parameter Selection for Derivative Estimation from Range Images" *Image and Vision Computing*, **13**, pp629–635, 1995..
22. M. Turner and E.R. Hancock, "Bayesian extraction of differential surface structure", *in Computer Analysis of Images and Patterns, Lecture Notes in Computer Science, Volume 970, Edited by V. Havlac and R. Sara*, pp. 784–789, 1995.
23. M. Turner and E.R. Hancock, "A Bayesian approach to 3D surface fitting and refinement", *Proceedings of the British Machine Vision Conference*, pp. 67–76, 1995.
24. G. Turk, "Re-tiling polygonal surfaces", *Computer Graphics*, **26**, pp. 55–64, 1992.
25. M. Vasilescu and D. Terzopoulos, "Adaptive meshes and shells", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* , pp. 829–832, 1992.
26. R. C. Wilson and E. R. Hancock, "Refining Surface Curvature with Relaxation Labeling", *Proceedings of ICIAP97*, Ed, A. Del Bimbo, Lecture Notes in Computer Science 1310, Springer pp. 150-157 1997.
27. R. C. Wilson and E. R. Hancock, "A Minimum-Variance Adaptive Surface Mesh", *CVPR'97*, pp. 634-639 1997.



(a) Rendering.



(b) Adaptive mesh.

(c) Normals.

(d) Curvature.

Fig. 4. Analysis of the range-data for the Mozart bust.