Scanning and Reconstruction of human body parts

N. A. Borghese and S. Ferrari

Laboratory of Human Motion Study and Virtual Reality, Istituto Neuroscienze e Bioimmagini CNR, via f.lli Cervi, 93, 20090 Segrate (Milano), Italy – borghese@inb.mi.cnr.it

Abstract

The reconstruction of 3D models of human body parts from range data is different with respect to that of general objects as they do not exhibit sharp discontinuities. Following this consideration, HRBF models which implement locally adapted filters are here introduced. They are based on stacking grids of Gaussians one over the other, where each grid operates at a different scale. The grids are not filled with Gaussians but these are inserted only in those crossings where the residual error is greater than the digitising noise. This allows to achieve a uniform reconstruction error. It results a very efficient and fast tool which can operate in real-time. Results on scanning a woman face are reported and discussed.

1. Introduction

3D scanning consists of fitting an analytical model to a set of 3D points sampled over the surface. In computer graphics, where the renderers work usually on triangles, the surface is represented as joint patches[1]. To obtain high definition and to preserve the exact topology (mainly vertexes and corners), this procedure becomes computationally expensive and cannot be carried out in real time. However, human body parts are, by their nature, relatively smooth and this property is exploited here to develop a more efficient procedure to recover their 3D shape.

We start from a set of 3D points randomly sampled over the surface with the only hypothesis that the surface has been heavily oversampled. Due to measurement noise, a simple triangulation (e.g. Delaunay tessellation) of the data points would produce an undesirable wobbling surface. To obtain a clean reconstruction, the approach proposed here is to use a reconstructor whose local frequency content can be automatically set from the data points. We observe that the spatial frequencies which constitute measurement noise are much higher than that of the body parts which, being smooth, feature low spatial frequencies. Therefore a low-pass filter with an adequate cut-off frequency will reconstruct the surface details and cut off the digitising noise. The Cut-off frequency is a very critical parameter: with a too low cut-off frequency, only the surface outline can be reconstructed and the most subtle details are left out. Increasing the cut-off frequency, the surface details are progressively outlined. However, the cut-off frequency cannot be increased ad libitum as a too high cut-off frequency yields the reconstruction of the original surface with the digitising noise imposed over it. A possibility is to increase the cut-off frequency until the following condition is met:

$$\frac{\int_D \left| S_a(P \mid \alpha) - z(P) \right| dP}{D} < \varepsilon \qquad (1)$$

where D is the input domain, z(P) are the points sampled over the surface and $S_a(P|\alpha)$ is the reconstructor. α are the parameters which define the cut-off frequency of S_a, and ε is related to the digitising error. A particularly suitable choice of the function $S_a(P|\alpha)$ is offered by the Hierarchical Radial Basis Function model (HRBF) proposed originally in the connectionist domain[3].

2. Method

$$S_r(P_k) = \sum_{l=1}^{L} \sum_{1}^{M_l} w_{kl} G(P_k \mid \sigma_l) \qquad (2)$$

In the HRBF model, the surface z(P) is reconstructed by adding the contributions of several grids of Gaussians, where each grid outputs a linear combination of its constituent Gaussians. The resulting surface, $S_r(P)$, is expressed as reported in Eq. (2) where L is the number of grids, M_l is the number of Gaussians in each grid, each centred in a grid crossing. w_{kl} is the weight associated to the k_{th} Gaussian of the l_{th} grid. All the Gaussians in one grid feature the same value of σ, σ_l, which determines the cut-off frequency of that grid, the smaller is σ_l, the higher is the cut-off frequency (the lower is the scale.

Grids with different σ_l allow the reconstruction of the surface at different scales. The first grid reconstructs the surface at a very coarse scale. Let us call $a_1(P)$ this reconstruction. For each data point a residual, $r_1(P_k)$, can be computed as: $r_1(P_k) = a_1(P_k) - z(P_k)$. In particular, for each Gaussian can be computed the local residual $r_1(k,l)$ as:

$$r_1(k,l) = \frac{\sum_{j=1}^{R_{k,l}} |r_j)|}{R_{k,l}} \tag{3}$$

where $R_{k,l}$ are all the points in the receptive field of the Gaussian, defined as the region $P_{k,l} \pm 2\sigma_l$.

$r_1(k,l)$ will be greater than noise in those regions where the details have not been reconstructed. This second grid will therefore be devoted to approximate this residual: its input will be $r_1(k,l)$ and its output the reconstruction, $a_2(P_k)$. A second residual, $r_2(P_k)$, is generated as: $r_2(P_k) = a_1(P_k) - r_1(P_k)$. With this second residual, a third grid is built and the procedure is repeated until the local residual error, goes under the noise threshold over all the input space (uniform convergence in L_1).

In the intermediate grids, the Gaussians will not be inserted in the whole space but only in those regions where the local residual in Eq. (3) is over the noise threshold. This allow to save many Gaussian units and computational time.

The weights, w_{kl} in Eq. (2), can be computed directly from the data points. In Borghese and Ferrari[2] it has been shown that the surface z(P) can be reliably reconstructed if we substitute the product of the surface height in the grid crossings times the square of the Gaussian spacing, Δx, to the weights:

$$S_r(P) = \sum_{l=1}^{L} \sum_{k=1}^{M_l} z(P_{kl}) G(P_{kl} \mid \sigma_l) \Delta x^2 \tag{4a}$$

$$z(P_{k,l}) = \frac{\sum_{j=1}^{R_{k,l}} z(P_r) G(P_r - P_{k,l} \mid \sigma_l)}{\sum_{1}^{R_{k,l}} G(P_r - P_{k,l} \mid \sigma_l)} \qquad P_r \in R \tag{4b}$$

The height of the surface is usually not available in the grid crossings or it is usually corrupted with noise. $z(P_{k,l})$ can be more advantageously determined through the local Maximum A-Posteriori Estimate, MAP of Eq. (4b).

3. Results and Conclusion

The reconstruction of a woman face is reported in Figure 1. 12,641 data points have been sampled over the face using the Autoscan system[4] as a digitiser with an accuracy of about 0.1mm. These points are fed to HRBF which furnishes a 3D model with a mesh size of 1.5mm. It should be remarked that the mesh size can be reduced ad libitum being $S_r(P)$, Eq. (4a), continuos. Colour and texture have been acquired from a photograph taken parallel to the face and mapped over the 3D model. The reconstruction requires only few seconds on a SGI Indigo2, 250Mhz processor, and can be implemented in real-time on a parallel machine.

Figure 1: 3D *Reconstruction of a woman's face.*

HRBF combines speed in the parameters computation with reconstruction accuracy and it is particularly suitable to be used with 3D digitisers to reconstruct in real-time a 3D surface from a set of sparse data points.

4. References

1. P. Cignoni, C. Montani, E. Puppo and R. Scopigno "Multiresoltion Representation and Visualization of Volume Data", *IEEE Trans. Visualization and Computer Graphics* **3**(4), pp. 352-369 (1997).
2. N.A. Borghese, S. Ferrari "Hierarchical RBF networks in function approximation" *Neurocomputing* **19**(1-3), pp. 259-283 (1998).
3. N.A. Borghese, G. Ferrigno, G. Baroni, R. Savarè, S. Ferrari and A. Pedotti "AUTOSCAN: A flexible and portable scanner of 3D surfaces", *IEEE Computer Graphics & Applications*, May/June, pp. 38-41 (1998).