Automatic Quasi-Isometric Surface Recovery and Registration from 4D Range Data

Toby Collins
Prof. Robert Fisher
University of Edinburgh

Dr. Adrien Bartoli
University of Clermont Ferrand

1 Introduction

To unlock the huge potential in video-rate range image analysis, arguably one the most significant hurdles to overcome is that of accurately recovering motion over a given sequence. For a single object, this largely amounts to uncovering the isomorphism undergone by the object's surface from a set of point sample observations in each frame. Unfortunately even with 3D data this is a highly challenging task, since in the absence of salient 3D structure (e.g. the region around a person's cheek or a section of cloth), motion cannot be resolved over the manifold (much like the aperture affect in 2D.) To alleviate this, we must impose assumptions about the object's deformation and/or appearance sub-space a priori (a concept well established in 2D image analysis). In 2D these priors must be extremely broad to account for the nonlinear, unknown and largely irrecoverable interaction between the object's deformed state and it's appearance in 2D. In general, the best we can do is to use extremely broad assumptions about smooth 2D motion, which has been demonstrated time again to be insufficient to adequately constrain the registration/tracking tasks.

By contrast, in our work we have focused on investigating general classes of deformation priors in 3D, where the story is very different. In 3D we can directly reason in terms of true 3D dynamics, and our assumptions on the underlying isomorphism can lead to very well defined registration problems. So well defined in fact that we show here how to go beyond mere registration (i.e. matching each frame to a particular template) to the harder problem of automatically recovering a complete, registered 3D model given only limited deformed observations. Solutions to this would find considerable application in the computer vision and graphics communities, where it is desirable to obtain object models that undergo non-rigid motion (e.g. a flying flag or a deforming body), undergoing occlusion (both self and external), never fully observable and present in cluttered scenes.

In this work our attention is focused on one important class of deforming objects; those whose surfaces undergo isometric or near-isometric deformations. These are a particular kind of isomorphism which relates a surface's embedding in 3D space (i.e. a Riemannian manifold) by a transformation preserving distances on the manifold (geodesics). Qausi-isometric transformations describe well the deformations undergone by a large range of real-world objects, such as many fabrics, paper, plastics, articulated objects and, to an extent, facial expressions. Robust methods for detecting and registering objects of this class therefore have uses in several fields including computer graphics, such as texture extraction, texture mapping, deformation transfer and video augmentation; computer vision, such as unsupervised deformable object learning, tracking, and deformation analysis.

2 Method Overview and Contribution

From a sequence of range observations our overarching goal is to automatically reconstruct the surfaces of deforming objects appearing in the sequence. This task is nontrivial for a number of reasons. Firstly, occlusion boundaries in range images, particularly those generated by stereo, cannot be reliably used for segmentation. Occlusion cues rely on depth discontinuities, which are often smoothed by the stereo algorithm, or not present if the occluding object is in contact with the surface. Without reliable segmentation, any priors on the isomorphism (including isometry or smoothness assumptions) are not valid (and thus detrimental) over regions which contain occlusion zones. The second problem that of global nonrigid alignment. Even if given a correct segmentation, matching pairs of range segments involves locally solving the non-rigid registration task. Without correspondences known a priori, this generally results in motion models with dense Jacobian and Hessian matrices. Scaling this up to a global alignment readily becomes intractable using numerical optimisation, yet this is highly desirable since local methods tend to result in global misalignments (as demonstrated in, for example 2D panorama stitching.)

However, we can obtain a foothold to the problem by considering certain deformation-invariant properties over the deforming surface. Here we present a global method for nonrigidly aligning surface segments located in range data that mutually agree with respect to the assumed isomorphism model. For near-isometry, this is the preservation of geodesic lengths on the manifold. We essentially perform surface model completion by embedded mosaicing, where a composite is formed in the surfaces intrinsic coordinate space. This has the desirable property that deformed segments are now relatable by far simpler transformations. For quasi-developable surfaces (i.e. isometric surfaces with very low Gaussian curvature such as cloth), isometry reduces to near-Euclidean transformations.

The key stages to our approach for automatic deformable surface reconstruction are shown in figure 1, and a brief overview of each stage is described in the following sections.

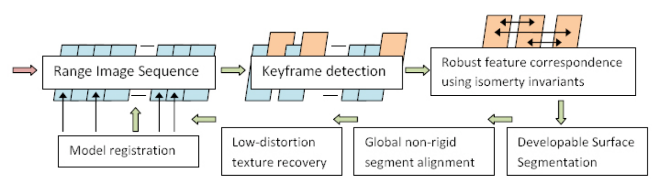

Figure 1: *Deformable model recovery and registration framework*

Since video rate range sequences contain a vast amount of redundant surface information, we would prefer to sample those frames which convey the most new surface information with which to build the model. Initially this set is selected using simple strategies (e.g. every n^{th} frame or random selection) from the range sequence, with further frame being added and included in the model based on (i) agreeing with the current model and (ii) revealing unseen sections of the surface.

2.1 Robust Feature Correspondence

Point correspondences are often key components in registration tasks largely for guiding the registration towards global optima. Typically, feature matchers in 2D (e.g. SIFT) and 3D (e.g. spin images) will misalign, so some form of correspondence annealing

Figure 2: *Deformable surface reconstruction from multiple range observations. TL: Regions matched and extracted from several range images (some of which are shown in TR. BL the resulting reconstruction comparing (left) the true surface texture, (middle) rigidly aligned patches and (right) the nonrigidly aligned and stitched model. BL: An example distortion map for a particular segment.*

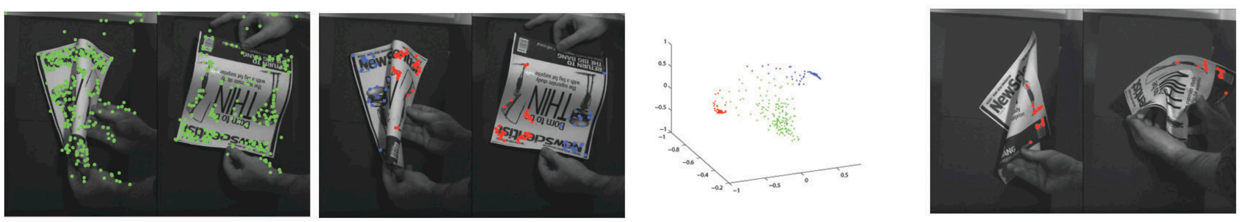

Figure 3: *Robust point correspondence between range pairs using spectral clustering. From left to right: Correspondences matched by SIFT; Robust correspondence and clustering using our method; Spectral embedding of matches (green indicates outliers); Another example.*

is often employed. We instead establish correspondence by appealing to the isometric assumption; that is mutual point distances on the manifold (geodesics) should be preserved. We establish a measure of isometric conformity between all pairs of candidate correspondences , and the optimal solution is the subset set of matches with maximal mutual agreement. In our method we find an efficient inexact solution to this using spectral methods to find strongly connected clusters in a compatibility graph whose nodes represent matches and edges weighted by their mutual match scores. Since occlusion zones (undetected or otherwise) break geodesic agreements, multiple clusters may exist in the graph which are each bounded by an occlusion zone. Thus, we use k-way spectral clustering to jointly recover the correct matches and segmentation (Figure 3.) Currently, we extract features based only on intensity, since 3D features on smooth developable surfaces are almost entirely ambiguous.

2.2 Quasi-Developable Surface Segmentation

Given a set of segmented high-quality feature correspondences, we proceed to extract the surrounding region agreeing with our assumption of developabliy. Although zero Gaussian curvature characterises perfectly developable surfaces, its computation is often too unstable to derrive a segmentation. Furthermore, for real surfaces low Gaussian curvature is usually present. Instead, from our matched features we grow a region over the range image such that the distortion (angular and stretch) induced by flattening it is bounded by some tolerance. Once grown, we further refine the region using other segmentation cues (i.e. strong intensity/depth gradients and colour histograms), combined proabablistically and solved using graph cuts.

2.3 Global Nonrigid Segment Alignment

Given multiple surface segments, from multiple range images, we bring them into nonrigid alignment by minimising two error criteria related to (i) mutual feature distances between matched segments and (ii) mutual feature distances over each segment. The first enforces the alignment of matched features whilst the second preserves the system's rigidity. Since (ii) is very nonlinear (quadric in position), we settle for a slightly weaker interpretation which enforces conformality rather than rigidity over a segment (i.e. angle preservation.) This is quadratic in position and results in a full-rank least squares system that can be solved in closed form using sparse linear least squares. The second stage propagates the transformations from the features through each segment. Since our underlying assumption is of near-rigidity on the 2D plane, we use an as-rigid-as-possible transformation similar to that proposed recently by [Schaefer et al. 2006] used in interactive shape manipulation.

2.4 Low-Distortion Surface Texture Recovery

Our final stage of recovering the surface model is to generate a rendering of the surface's texture. Given the match correspondences, shading artifacts can be removed using standard techniques. To generate the render, blending techniques used in image mosaicing are not suitable, due to residual ghosting from small misalignments. Instead we are driven by the goal of a low-distortion rendering. In our method we greedily select and stitch those regions which were were transformed onto the plane with the least distortion. Given an initial patch (i.e. the one least distorted), we treat the addition of a second patch as a binary labeling problem, combining a distortion penalty with a seam cost, and solve this incrementally using graph cuts.

2.5 Registration

In order to fit the model to the rest of the data, we adopt an energy minimisation process which penalises states that diagree in data evidence (i.e. 3D distance to the point cloud and feature correspondence), and a quadratic isometric within-plane bending model [Bergou et al. 2006]. For all other frames, we perform the same strategy but initialise the model to its deformed state in the previous frame. Global drift is avoided by using hard feature constraints between the model and scan as described in section 2.1. The newly registered instances can then be further incorporated into our model, along with new neighbouring regions agree with respect to the model's isomorphism prior.

3 Results

We have tested our approach on several sequences capturing deforming paper and fabrics. Figures 2 and 3 show examples of some good results attained on a seqeunce of a deforming magazine cover. In the near future we aim to further validate our methods for other sequences. We also aim to extend our work by better modelling the relationship between mesh distortion and nonrigid deformation constraints on the 2D plane, and will be investigating our framework for recovering other classes of deforming surfaces.

References

BERGOU, M., WARDETZKY, M., HARMON, D., ZORIN, D., AND GRINSPUN, E. 2006. A quadratic bending model for inextensible surfaces. In *SGP '06: Proceedings of the fourth Eurographics symposium on Geometry processing*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 227–230.

SCHAEFER, S., MCPHAIL, T., AND WARREN, J. 2006. Image deformation using moving least squares. In *International Conference on Computer Graphics and Interactive Techniques*.