Global temporal registration of multiple non-rigid surface sequences

Peng Huang, Chris Budd, Adrian Hilton

Centre for Vision Speech Signal Processing, University of Surrey
University of Surrey, Guildford, UK, GU2 7XH

{p.huang, c.budd, a.hilton}@surrey.ac.uk

Abstract

In this paper we consider the problem of aligning multiple non-rigid surface mesh sequences into a single temporally consistent representation of the shape and motion. A global alignment graph structure is introduced which uses shape similarity to identify frames for inter-sequence registration. Graph optimisation is performed to minimise the total non-rigid deformation required to register the input sequences into a common structure. The resulting global alignment ensures that all input sequences are resampled with a common mesh structure which preserves the shape and temporal correspondence. Results demonstrate temporally consistent representation of several public databases of mesh sequences for multiple people performing a variety of motions with loose clothing and hair.

1. Introduction

Non-rigid surface reconstruction from single and multiple view video has advanced to allow reconstruction of detailed surface deformation such as cloth. For scenes with complex non-rigid deformation multiple-view reconstruction is commonly performed independently at each frame of the sequence resulting in the reconstruction of an unstructured sequence of meshes with different vertex positions and connectivity. Recent research has focused on the problem of extracting a temporally consistent mesh representations across sequences using both model-based [1] and model-free [3, 6, 17, 19, 22] approaches. Single sequence registration approaches commonly employ a sequential frame-to-frame registration approach assuming relatively small non-rigid deformation between consecutive frames. Impressive results have been demonstrated on highly non-rigid surfaces such as people with loose clothing.

In this paper we address the problem of obtaining a consistent mesh representation across multiple sequences of the same non-rigid object performing different motions. This case commonly occurs for example in capturing a

database of human movements composed of different short-clips [10, 19]. Global alignment of the mesh sequences into a consistent mesh structure is important for the analysis of non-rigid surface dynamics, efficient storage and reuse of the mesh sequences in computer graphics applications.

A global non-rigid surface alignment framework is presented which minimises the deformation required to register all meshes from different input sequences into a temporally consistent structure. Minimisation of non-rigid deformation is based on the shape similarity between meshes. Shape similarity is first employed to identify frames between different sequences with similar shape which can be registered building on existing frame-to-frame non-rigid alignment techniques [1, 4, 16]. A global graph path optimisation is then performed to minimise the total non-rigid deformation and select the best reference frame for resampling to a temporally consistent mesh structure. This results in a temporally consistent mesh structure and correspondence across all sequences. Results demonstrate temporally coherent mesh sequences obtained for several public databases of multiple people performing a variety of motions including complex break dance sequences and loose clothing. Quantitative evaluation demonstrates that the resulting temporally aligned sequences give an accurate approximation of the unaligned input sequences.

2. Related Work

Non-rigid surface reconstruction from single or multiple view video commonly recovers an independent surface reconstruction at each time instant which lacks the temporally coherent structure and motion of the observed scene. Surface structure and motion estimation has been addressed using flow-based algorithms to track surface in multiple view video sequences. Vedula et al. [20] introduced the concept of scene-flow as the extension of 2D optic-flow to 3D deforming surfaces. Carceroni and Kutulakos [7] compute shape and non-rigid motion under known illumination using relatively large 3D surface elements. Pons et al. [13] introduce a variational approach providing a global solution for scene flow. Cagniart et al. [4, 5, 6] introduced a local

Figure 1. Overview flow diagram. Left to right: Multiple view video capture on n -different sequences (typically 30-120s) are first individually reconstructed; then temporal alignment within each sequence; global temporal alignment combining multiple temporally aligned sequences into a single temporally consistent mesh.

rigid patch based representation for tracking non-rigid surfaces to obtain a single temporally coherent structure and scene flow estimation. This approach achieves consistent structure and motion estimation for sequences of complex human movement including multiple people and objects.

Non-rigid surface tracking using prior models of surface shape and appearance has been widely used for both sequential tracking and registration across wide-timeframes. Deformable surface models have been widely used in the medical domain to recover anatomical structure and motion from 3D image sequences based on landmark correspondence [11]. Salzmann et al. [14] construct a parameterised deformable model for inextensible manifold surfaces which is used to recover 3D surface shape from monocular image sequences. Furukawa et al. [8] introduce a joint formulation of multiple view reconstruction and tracking using the initial frame as a reference deformable model to recover non-rigid 3D facial and cloth movement. Articulated deformable surface models have been employed to constrain the estimation of human surface motion from single and multi-view video [12]. To recover more detailed non-rigid motion of people and clothing from multiple view reconstructions Aguiar et al. [1] employ an initial high-resolution 3D body scan and use an implicit volumetric parameterisation of the shape to constrain the deformation. Recent approaches [15, 2, 9] have employed learnt statistical models of human shape and pose to recover both non-rigid shape and motion from monocular image sequences. Stoll et al. [18] combine non-rigid surface tracking and a high-resolution 3D body scan with statistical body models and physics-based cloth simulation to estimate and reproduce highly realistic non-

rigid cloth dynamics. Model-based approaches are generally restricted to a specific class of non-rigid objects such as people or faces and may not correctly align non-rigid deformations which are not represented.

Scene flow approaches estimate pairwise frame-to-frame temporal alignment which may fail if the non-rigid deformation between frames is too large to locally estimate the dense correspondence. Some research has investigated the problem of establishing dense registration of surfaces subject to large non-rigid deformations. Bronstein et al. [3] use generalised multidimensional scaling to construct a surface parameterisation which preserves geodesic distance under isometric non-rigid deformation. Starck et al. [17] proposed a multi-resolution belief-propagation framework incorporating both photometric and geometric features to establish dense correspondence between non-rigid surfaces for wide-timeframes or different sequences. Tung et al. [19] propose a geodesic surface matching approach to estimate correspondence across both wide-timeframes and between different non-rigid objects. Zeng et al. [22] propose a graph matching framework for dense non-rigid surface registration of objects in widely varying poses. These approaches provide a method of establishing dense correspondence between individual frames of non-rigid surface sequences captured at different times. However, the computational cost of the approaches is relatively high compared to frame-to-frame scene flow alignment prohibiting their use for dense correspondence across sequences.

In this paper we investigate the problem of dense registration for a database containing multiple non-rigid surface sequences of multiple motions of the same object captured

at different times into a single temporally coherent structure. This builds on existing frame-to-frame surface tracking methods to obtain an independent temporal structure for each sequence. The problem is then to register these sequences into a representation which is consistent across sequences.

3. Global Non-rigid Alignment

We can state the problem of global non-rigid alignment of multiple temporally aligned mesh sequences as follows. Given a set of temporally aligned non-rigid mesh sequences $\{S_i\}_{i=1}^N$ of a deforming object such that each mesh sequence S_i is a temporally aligned sequence of meshes $S_i = \{M_i(t)\}_{t=0}^{T_i}$ where mesh $M_i(t) = (C_i, X_i(t))$ has constant vertex connectivity C_i and time varying vertex positions $X_i(t)$, then the problem is to obtain a single temporally consistent mesh representation for all sequences such that for each frame $\hat{M}_i(t) = (\hat{C}, \hat{X}_i(t))$ where $\hat{C}$ is the global connectivity which is the same for all meshes $M_i(t)$ for all sequences S_i . In addition we require that two conditions are satisfied: (1) shape is preserved such that all vertices for the temporally aligned mesh lie on the unaligned mesh $\hat{X}_i(t) \in M_i(t)$ and vice-versa $X_i(t) \in \hat{M}_i(t)$; and (2) temporal alignment of individual sequences is preserved which can be stated as follows, if k^{th} vertex of an aligned mesh $\hat{x}_{ik}(t) \in \hat{X}_i(t)$ lies at position $p(t) \in M_i(t)$ then we want $\hat{x}_{ik}(t+1) = p(t+1)$.

Our approach to global alignment comprises 4 steps:

- (1) Inter-sequence similarity: identify frames in each pair of sequences $M_i(t_r) \in S_i$ and $M_j(t_s) \in S_j$ with similar shape.
- (2) Inter-sequence frame registration: align meshes with similar shape $M_i(t_r)$ and $M_j(t_s)$ using existing frame-to-frame single sequence non-rigid alignment techniques to obtain $\hat{M}_j(t_s)$ with the same mesh connectivity as $M_i(t_r)$.
- (3) Global alignment graph construction: construct a graph representing all the pair wise inter-sequence frame registrations.
- (4) Global non-rigid alignment: All frames for all sequences are aligned and remeshed with a single mesh connectivity $\hat{C}$ using the shortest-path through the alignment graph.

The global alignment graph for a small set of four sequences is illustrated in figure 2. This approach assumes that at least one sufficiently similar frame exists between sequence S_i and one of the other sequences S_j for $j \neq i$, where *sufficiently similar* is determined by the ability of the selected single sequence alignment method to obtain a valid frame-to-frame registration which satisfies conditions (1) and (2) in the problem statement. If this is not satisfied then wide-timeframe alignment [3, 17, 19, 22] could be used to automatically register frames with dissimilar non-rigid shapes at the expense of increased computational cost. In

practice for databases comprising reconstructed non-rigid sequences of people performing a variety of movements use of existing frame-to-frame alignment provides sufficient inter-sequence registrations for global alignment. In the remainder of this section we present the details of the global alignment approach starting with an outline of the single sequence alignment approach which is used in step (2) for inter-sequence registration of single frames.

3.1. Single Sequence Alignment

The global alignment framework introduced in this paper assumes that the input mesh sequences have been independently temporally aligned. In this work we extend the local rigid patch based sequence alignment approach [5] by incorporating photometric features to reduce drift across the surface and allow larger frame-to-frame deformations. This approach is used in this paper to register individual *transition* frames between sequences of different motions.

For a sequence of meshes $\{M_i(t_r)\}$ temporal consistency is created by sequential frame-to-frame alignment. Slightly different from [5] which estimates rigid transformations for patches and yields dense point trajectories, sparse correspondence between consecutive frames $M_i(t_r)$ and $M_i(t_{r+1})$ is estimated between patches combined with photometric SIFT features to reduce drift. These feature matches provide constraints for a coarse-to-fine multi-resolution deformation of the mesh $M_i(t_r)$ using a volume-preserving Laplacian deformation approach [16, 1] based on $\delta = Lv_0$, where δ are the differential coordinates of a tetrahedralization of the mesh, L is the Laplacian operator and v_0 are the mesh vertex positions. The surface of mesh $M_i(t_r)$ is regularly divided into m patches based on geodesic distance allowing selection of the best subset of evenly distributed features. To compute new vertex locations for deformed mesh $\hat{M}_i(t_r)$ within this system, we minimise the combined error:

$$\arg \min_v \{ \|Lv - \delta\| + \|W_c(v - v_c)\| \} \quad (1)$$

where v is the set of deformed vertex positions, v_c is the set of known vertex positions from feature constraints, W_c is a diagonal weight matrix with 1 in positions corresponding to constrained vertices and 0 otherwise. The first and second term in equation 1 correspond to rigidity and smoothness costs respectively [16] with the feature matches as soft constraints. Multi-resolution solution with increasing $m \in [30, 150]$ increases the range of convergence for the geometric rigid patch matching allowing alignment of sequences with large non-rigid deformation for input to the global inter-sequence alignment and inter-sequence frame alignment.

Figure 2. Illustration of a global alignment graph for four sequences showing edges for inter-sequence frame registration and successive frames.

3.2. Inter-sequence similarity

The basis for our approach to global alignment of multiple non-rigid surface sequences is the identification of *similar* frames between different sequences. Similar frames in different sequences are identified using a non-rigid shape similarity metric which does not require known surface correspondence. Similar frames between different sequences are then used for non-rigid alignment of a single frame using the approach presented in section 3.1. This provides an inter-sequence registration at multiple sparse frames throughout the databases of non-rigid surface sequences.

Given two temporally aligned mesh sequences S_i and S_j the inter-sequence shape similarity $c(M_i(t_r), M_j(t_s))$ between the meshes two frames $M_i(t_r)$ and $M_j(t_s)$ is ideally defined based on the correspondence. However, prior to global alignment the correspondence is unknown, we therefore define an approximate shape similarity measure. Shape histograms have previously been demonstrated to give good performance for measuring non-rigid deformable shape similarity [10]. In this work we use a volumetric shape histogram where the space is sub-divided into radial and angular bins located at the centroid of the shape, $(\Delta r, \Delta \theta, \Delta \phi)$. The volumetric shape histogram $H(M)$ represents the spatial occupancy of the bins for a given mesh M . We can then define a measure of shape dissimilarity between two mesh $M_i(t_r)$ and $M_j(t_s)$ by optimising for the maximum overlap between their corresponding radial bins with respect to rotation about the centroid.

$$c(M_i(t_r), M_j(t_s)) = \min_{\theta, \phi} \|H(M_i(t_r)) - H(M_j(t_s), \theta, \phi)\| \quad (2)$$

where $\|\cdot\|$ denotes the sum of squared differences over all bins in the histogram and $H(M, \theta, \phi)$ is the spherical histogram $H(M)$ rotated by angles (θ, ϕ) for invariance to orientation. For computational efficiency the similarity matrix between all frames in all sequences is evaluated once for subsequent use. Example similarity matrices for multiple non-rigid mesh sequences of multiple person are shown in Figure 4.

3.3. Inter-sequence frame registration

Inter-sequence frame registration for sequences S_i and S_j is performed on the frames which minimise the shape dissimilarity measure $c()$ defined in equation 2. The optimal frames for registration are defined as:

$$(r_{opt}, s_{opt}) = \arg \min_{r, s} c(M_i(t_r), M_j(t_s)) \quad (3)$$

Registration of $M_i(t_{r_{opt}})$ and $M_j(t_{s_{opt}})$ is performed using the single sequence alignment approach presented in section 3.1 giving a non-rigidly deformed source mesh $\tilde{M}_j(t_s)$ which has the same connectivity C_i as mesh M_i from the i^{th} sequence and the same shape as the original target mesh $M_j(t_s)$.

To evaluate if the shape preservation condition (1) for global mesh alignment is satisfied we can compute the Eu-

clidean distance between every vertex in the deformed-source-to-target mesh and vice versa. This approach is computationally expensive due to the requirement for $O(N_v^2)$ nearest point computations, where N_v is the number of mesh vertices. For efficiently we equivalently evaluate the shape similarity $c(\hat{M}_j(t_s), M_j(t_s))$ between the deformed-source and target mesh to determine if the shape alignment has performed correctly according to a fixed similarity threshold representing a 95% overlap in shape. If the alignment does not satisfy this shape similarity condition we reject the corresponding inter-sequence registration.

3.4. Global Alignment Graph Construction

An alignment graph is constructed which links the set of input sequences $\{S_i\}$ together with nodes representing the individual meshes $M_i(t)$ for each frame and edges representing inter-sequence frame transitions with valid registered meshes as identified in section 3.2. The alignment graph for a small set of sequences is illustrated in figure 2. This graph structure provides the basis for aligning the set of input sequences to obtain a representation with a single mesh connectivity $\hat{C}$ across all input frames.

The graph defines the set of possible paths connecting frames between any two input meshes for $M_i(t_r)$ and $M_j(t_s)$. A path $P_{irjs} = \{M_q\}_{q=0}^{N_p}$ through the graph is a set of successive frames from the input sequences and inter-frame transitions such that $M_0 = M_i(t_r)$ and $M_{N_p} = M_j(t_s)$. The cost $C_P(M_i(t_r), M_j(t_s))$ of alignment between a pair of input frames for a particular graph path P_{irjs} can then be defined as the sum of the dissimilarities between successive mesh frames along the path:

$$C_P(M_i(t_r), M_j(t_s)) = \sum_{q=0}^{N_p-1} c(M_q, M_{q+1}) \quad (4)$$

Optimal alignment of any two input frames will be given by the shortest graph path $P_{opt}(i, r, j, s)$ which minimises equation 4 as this minimises the non-rigid deformation which must be applied to successive meshes along the path for alignment to a temporally consistent mesh structure with the same mesh connectivity. The shortest path can be evaluated without knowledge of the correspondence between meshes along the path due to the use of the shape similarity metric to define the cost.

3.5. Global Non-rigid Alignment

Global alignment is performed by optimising the shortest path between input frames to minimise the total non-rigid deformation. Selecting as the reference a single frame M_{ref} we evaluate the shortest path to all other frames in the dataset. The optimal reference mesh M_{ref} and shortest

paths which minimise the total non-rigid deformation for alignment to all other frames is then evaluated as:

$$C_{total} = \min_{M_{ref}, P_{it}} \sum_{i=1}^N \sum_{t=0}^{T_i} C_P(M_i(t), M_{ref}) \quad (5)$$

As the inter-frame registrations between sequences are sparse the computation can be performed efficiently by propagating optimal path costs for each frame across the graph from M_{ref} . Giving the set of shortest graph paths for all frames $P_{opt}(it)$ which minimise the non-rigid deformation from the reference frame. This operation integrates the similarity $c()$ for successive frames in a sequence and across inter-sequence transitions. C_{total} is a measure of the total non-rigid deformation for global alignment of all frames with a temporally consistent mesh structure.

Having defined the shortest path from M_{ref} for each frame we can resample the mesh for every frame to have the same globally consistent mesh connectivity as the reference frame $\hat{C} = C_{ref}$. The vertex positions for the mesh $\hat{M}_i(t_r)$ at each frame are obtained by resampling the original mesh $M_i(t_r)$. Details of the vertex resampling for remeshing are presented in Appendix A.

Global non-rigid alignment gives a mesh for each frame $\hat{M}_i(t_r) = (\hat{C}, \hat{X}_i(t_r))$ which has a constant global connectivity $\hat{C}$, preserves the shape of the original mesh and preserves the temporal alignment of the input non-rigid mesh sequences and computed inter-sequence frame alignments. The graph path optimisation defined by equation 5 minimises the total non-rigid deformation required to align all sequences.

4. Experimental Results and Evaluation

Global alignment has been applied to four databases of 3D mesh sequences for people performing a variety of motions: MIT [21] and Street Dancer, Character1, Fashion1[17]. The number of frames for each sequence varies from 30-500, details of the number of frames for each sequences together with the total number of frames are shown in brackets in Table 1. Single sequence temporal alignment was initially performed on each sequence using existing methods, section 3.1. Shape similarity is computed with a spherical shape histogram of 1.5m radius with bin sizes $(\Delta r, \Delta \theta, \Delta \phi) = (0.3m, 18^\circ, 18^\circ)$. The shape similarity matrix for each database is presented in Figure 4 where the normalised colour map from dark blue to red indicates high to low similarity. The structure of the similarity matrix shows that there are large differences between frames due to the non-rigid surface deformation.

Global registration is performed using the approach presented in section 3 to obtain a single temporally consistent mesh for each sequence. Computation time for the

(a) Street Dancer

(b) MIT

(c) Character1

(d) Fashion1

Figure 3. Example frames of resulting consistent mesh per database (top to bottom: StreetDancer, MIT, Character, Fashion).

graph optimisation is $< 2s$ total and resampling takes approximately $10s/frame$. Example frames from different sequences for each database are shown in Figure 3. The surfaces are textured with the same pattern across all sequences based on the global non-rigid alignment. The results demonstrate consistent placement of the texture pattern across all sequences. The accompanying video shows the original sequences and globally aligned texture map re-

sults. The stability of the texture map placement on the deforming surface indicates accurate alignment between sequences. Quantitative results for the global representation accuracy are presented in Table 1. This shows the worst-case results computed using the Hausdorff distance between the original and resampled mesh as minimum and maximum errors across all meshes in each sequence. Sequences with zero error indicate the sequence from which the op-

Database	Sequences	Min. (cm)	Max. (cm)	RMSE (cm)
StreetDancer (1664)	flashkick (199)	3.66	4.72	0.20
	free (499)	2.37	2.77	0.05
	head (249)	2.54	3.69	0.11
	kickup (219)	0.00	0.00	0.00
	lock (249)	2.51	3.55	0.08
	pop (249)	2.03	2.64	0.10
MIT (347)	bouncing2 (174)	0.00	0.00	0.00
	handstand2 (173)	3.79	4.95	0.19
Character1 (442)	Hit (76)	2.68	4.36	0.51
	Jog (26)	2.66	3.46	0.15
	Jog2Walk (28)	3.91	6.18	0.62
	Stagger (81)	2.25	3.94	0.55
	Stand (50)	2.84	4.14	0.26
	Stand2Walk (31)	1.98	2.78	0.21
	Tense (36)	2.55	3.12	0.14
	Walk (31)	2.47	3.42	0.27
	Walk2Jog (27)	3.32	4.23	0.27
	Walk2Stand (56)	0.00	0.00	0.00
Fashion1 (371)	Pose (125)	2.94	3.43	0.11
	Stand2Walk (48)	0.00	0.00	0.00
	Twirl (40)	4.27	5.25	0.35
	Walk (30)	3.15	4.91	0.41
	Walk2Stand (56)	2.67	3.92	0.32
	WalkPose (72)	3.23	5.55	0.58

Table 1. Evaluation of worst-case Hausdorff distance shape error for global alignment over all mesh sequences

timal reference frame was selected. In the case of Fashion1’s Twirl, the alignment error is relatively large. This can be explained as the poor non-rigid alignment of the loose skirt material during the foot rotational motion. All sequences show a root mean square error (RMSE) of alignment meshes $< 1\text{cm}$ indicating good overall representation of the captured 3D shape.

5. Conclusion

We have presented a novel framework for global alignment of multiple surface sequences of a non-rigid object. The approach builds on existing approaches for frame-to-frame alignment across a single sequence to enable construction of a temporally consistent representation for multiple sequences of different non-rigid motions. A global alignment graph is introduced based on shape similarity to represent the sparse inter-sequence registration of frames. Graph path optimisation is used to determine the global sequence alignment and reference frame which minimise the non-rigid mesh deformation for representation of all frames with the same temporally consistent mesh structure. This results in a temporally consistent representation with the

Figure 4. Shape similarity matrix for a set of sequences of a person performing various motions, colour map shows (blue) maximum similarity to (red) maximum dissimilarity

same mesh connectivity at all frames, which preserves both the shape and non-rigid correspondence of the input sequences.

Global alignment is applied to four public databases of multiple mesh sequences for different people performing a variety of movements. Results and quantitative evaluation demonstrate that the approach results in a single temporally coherent mesh structure with accurate representation of the original shape and non-rigid motion. This multiple sequence non-rigid surface alignment is an important step for subsequent analysis, compression and manipulation or mesh sequences reconstructed from multiple view video. This approach uses shape similarity and is therefore not sensitive to topology change and will work for different characters if they have similar shape. To the best of our knowledge this is the first approach to recover temporally consistent structure across a database of multiple non-rigid mesh sequences.

Acknowledgement

This work was funded by EU IST project i3DPost <http://www.i3dpost.eu> Many thanks to Cedric Cagniart and Edmond Boyer for the access to non-rigid alignment of StreetDancer and MIT sequences.

Appendix A: Mesh Resampling

Given two meshes X and Y with different mesh structure but in the same shape, the geodesic relationship between a vertex $\mathbf{v}_p \in X$ and a closest triangle face $\mathbf{f}_{\text{opt}} \in Y$ is computed and stored,

$$\mathbf{f}_{\text{opt}} = \arg \min_{f \in Y} \sum_{i=0}^2 |\mathbf{v}_p - \mathbf{v}_i(\mathbf{f})| \quad (6)$$

The relationship between $\mathbf{v}_p$ and the corresponding $\mathbf{f}_{\text{opt}} = \langle \mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2 \rangle$ is presented as follows,

$$\mathbf{v}_p = \alpha \mathbf{v}_0 + \beta \mathbf{v}_1 + \gamma \mathbf{v}_2 \quad (7)$$

This linear equation is solved to obtain (α, β, γ) . For an unprocessed mesh $\hat{Y}$ which has the same mesh structure as Y but a different shape. We can then compute the corresponding point $\mathbf{v}_q$ to $\mathbf{v}_p$ by following,

$$\mathbf{v}_q = \alpha \mathbf{u}_0 + \beta \mathbf{u}_1 + \gamma \mathbf{u}_2 \quad (8)$$

where $(\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2) \in \hat{Y}$ correspondes to $(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2) \in Y$ respectively. Finally, a new mesh $\hat{X}$ which has the same mesh structure as X and the same shape as $\hat{Y}$ is obtained by updating each $\mathbf{v}_p \in X$ to the position of vertex $\mathbf{v}_q$.

References

- [1] E. Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H.-P. Seidel, and S. Thrun. Performance Capture from Sparse Multi-view Video. *Proc. ACM SIGGRAPH*, 27(3), 2008. 1, 2, 3
- [2] A. Balan, M. Black, J. Davis, and H. Haussecker. Detailed Human Shape and Pose from Images. In *Conference on Computer Vision and Pattern Recognition*, 2010. 2
- [3] A. Bronstein, M. Bronstein, and R. Kimmel. Calculus of non-rigid surfaces for geometry and texture manipulation. *IEEE Trans. Visualization and Computer Graphics*, 13(5):902–913, 2007. 1, 2, 3
- [4] C. Cagniart, E. Boyer, and S. Ilic. Free-Form Mesh Tracking: a Patch-Based Approach. In *Conference on Computer Vision and Pattern Recognition*, 2010. 1
- [5] C. Cagniart, E. Boyer, and S. Ilic. Iterative deformable surface tracking in multi-view setups. In *3DPVT*, 2010. 1, 3
- [6] C. Cagniart, E. Boyer, and S. Ilic. Probabilistic Deformable Surface Tracking from Multiple Videos. In *European Conference on Computer Vision*, 2010. 1
- [7] R. Carceroni and K. Kutulakos. Multi-view scene capture by surfel sampling: From video streams to non-rigid 3d motion, shape and reflectance. *International Journal of Computer Vision*, 49(2-3):175–214, 2002. 1
- [8] Y. Furukawa and J. Ponce. Dense 3D Motion Capture for Human Faces. In *Conference on Computer Vision and Pattern Recognition*, 2010. 2
- [9] H. Hasler, N. Ackermann, B. Rosenhahn, T. Thormahlen, and H.-P. Seidel. Multi-linear Pose and Body Shape Estimation of Dressed Subjects from Image Sets. In *Conference on Computer Vision and Pattern Recognition*, 2010. 2
- [10] P. Huang, A. Hilton, and J. Starck. Shape Similarity for 3D Video Sequences of People. *International Journal of Computer Vision*, 89(2-3):362–381, 2010. 1, 4
- [11] T. McInerney and D. Terzopoulos. Deformable models in medical images analysis: a survey. *Medical Imaging Analysis*, 1(2):91–108, 1996. 2
- [12] T. Moeslund, A. Hilton, and V. Kruger. A Survey of Advances in Vision-Based Human Motion Capture and Analysis. *Computer Vision and Image Understanding*, 104(2-3):90–127, 2006. 2
- [13] J.-P. Pons, R. Keriven, and O. Faugeras. Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score. *International Journal of Computer Vision*, 72(2):179–193, 2007. 1
- [14] M. Salzmann, J. Pilet, S. Ilic, and P. Fua. Surface deformation models for non-rigid 3-d shape recovery. *to appear IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007. 2
- [15] L. Sigal, A. Balan, and M. Black. Combined discriminative and generative articulated pose and non-rigid shape estimation. In *Conference on Computer Vision and Pattern Recognition*, 2010. 2
- [16] O. Sorkine, D. Cohen-Or, Y. Lipman, M. Alexa, C. Rössl, and H.-P. Seidel. Laplacian surface editing. In *SGP '04: Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 175–184, New York, NY, USA, 2004. ACM. 1, 3
- [17] J. Starck and A. Hilton. Correspondence labelling for wide-timeline free-form surface matching. In *IEEE Int.Conf.on Computer Vision*, 2007. 1, 2, 3, 5
- [18] C. Stoll, J. Gall, E. Aguiar, S. Thrun, and C. Theobalt. Video-based Reconstruction of Animatable Human Characters. In *ACM SIGGRAPH ASIA*, 2010. 2
- [19] T. Tung and T. Matsuyama. Dynamic Surface Matching by Geodesic Mapping for Animation Transfer. In *Conference on Computer Vision and Pattern Recognition*, 2010. 1, 2, 3
- [20] S. Vedula, S. Baker, P. Rander, R. Collins, and T. Kanade. Three-dimensional scene flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):475–480, 2005. 1
- [21] D. Vlasic, I. Baran, W. Matusik, and J. Popović. Articulated mesh animation from multi-view silhouettes. In *Proc. ACM SIGGRAPH*, pages 1–9, 2008. 5
- [22] Y. Zeng, C. Wang, Y. Wang, X. Gu, D. Samaras, and N. Paragios. Dense Non-rigid Surface Registration Using High-Order Graph Matching. In *Conference on Computer Vision and Pattern Recognition*, 2010. 1, 2, 3