Model-based Integration of Visual Cues for Hand Tracking

Shan Lu
CS Department
Rutgers University
Piscataway, NJ 08854

Gang Huang
CIS Department
University of Pennsylvania
Philadelphia, PA 19104

Dimitris Samaras
CS Department
S.U.N.Y. Stony Brook
Stony Brook, NY 11794

Dimitris Metaxas
CS Department
Rutgers University
Piscataway, NJ 08854

Abstract

We present a model based approach to the integration of multiple cues for tracking high degree of freedom articulated motions. We then apply it to the problem of hand tracking using a single camera sequence. Hand tracking is particularly challenging because of occlusions, shading variations, and the high dimensionality of the motion. The novelty of our approach is in the combination of multiple sources of information which come from edges, optical flow and shading information. In particular we introduce in deformable model theory a generalized version of the gradient-based optical flow constraint, that includes shading flow i.e., the variation of the shading of the object as it rotates with respect to the light source. This constraint unifies the shading and the optical flow constraints (it simplifies to each one of them, when the other is not present). Our use of cue information from the entirety of the hand enables us to track its complex articulated motion in the presence of shading changes. Given the model-based formulation we use shading when the optical flow constraint is violated due to significant shading changes in a region. We use a forward recursive dynamic model to track the motion in response to 3D data derived forces applied to the model. The hand is modeled as a base link (palm) with five linked chains (fingers) while the allowable motion of the fingers is controlled by recursive dynamics constraints. Model driving forces are generated from edges, optical flow and shading. The effectiveness of our approach is demonstrated with experiments on a number of different hand motions with shading changes, rotations and occlusions of significant parts of the hand.

1. Introduction

In this paper we present a model based approach to high degree of freedom articulated motion tracking, based on the integration of visual cues and apply it to the problem of hand tracking using a single camera sequence. Hand tracking has received significant attention in the last few years, because of its crucial role in the design of new human computer interaction methods, gesture analysis and sign language understanding. Glove based devices capture human hand motion directly, but are expensive and hard to use. Vision-based hand tracking is a cost-effective, non invasive alternative. Serious challenges lie in the high number of degrees of freedom and the problem of occlusions.

Two general approaches have been suggested for this problem. Model based approaches try to estimate the position of a hand by projecting a 3-D hand model to image space and comparing it with image features (fingertips [26, 25, 27], line segments [26]). A spline based hand shape model was used in [24] to minimize differences between the silhouettes. Others [30, 26] have used stereo to avoid occlusions. Appearance based approaches estimate hand postures directly from the images after learning the mapping from image feature space to hand configuration space [29, 28]. Such systems are more useful for recognizing discrete hand states than for general purpose hand tracking.

Study of motion and shading together has been recently formalized [21, 23] and extended to multiple views [22]. Our approach is model-based and hence can work with a single view. Our first contribution is in the combination of cue forces from edges, optical flow and shading. In particular we introduce in deformable model theory a generalized version of the gradient-based optical flow constraint, that includes shading flow i.e., the variation of the shading of the object as it rotates with respect to the light source. This constraint unifies the shading and the optical flow constraints and degenerates to each one of them when the other is not present. Although optical flow and edges in deformable models have been used in the past [20], as well as shading [19], these two methods were applied to different problem domains (moving and static objects respectively). In this paper we combine them to correct for the errors due to the brightness constancy assumption. We use cue information from the entirety of the hand and we are able to track its complex articulated motion in the presence of shading changes. Given the model-based formulation we augment the optical flow constraint with shading information.

The hand can have as many as 26 degrees of freedom when we model it as a multiple open chain structure. The dynamic/kinematic problem of such a large system, which contains not only open chains but also closed chains, can be modeled as a sub-problem of robotic mechanisms. There are many forward and inverse dynamics simulation techniques for human and robotic motion [14], [16], [18] [10], [17], [13], [15]. The second contribution of our approach is the use of a forward recursive dynamic model, to track motion in response to 3D data derived forces applied to the model. The hand is modeled as a base link (palm) with five linked chains (fingers). Using such a formulation we limit the allowable motion of the fingers with the use of recursive dynamics constraints. The model's driving forces are computed from image cues such as edges, optical flow and shading.

In our formulation we compute from edge, optical flow and shading cue constraints 2D data-based forces. The perspective camera model is used to convert these 2D forces into 3D forces that drive the hand model. These 3D forces are then used to calculate the acceleration of our dynamic hand, its new velocity and new position. Since this is a second order dynamic hand model we use it to predict finger motion from one frame to the next so that we are closer to the data in the next frame. To avoid unnecessary calculations of the shading constraint we monitor the intensity changes in several hand areas during tracking and use it only if these changes are significant.

The dynamic hand model is described in Sec. 2. Sec. 3 presents model initialization and generation of image forces. Sec. 4 introduces illumination information on the optical flow constraint. Recursive dynamics of the hand model and constraints on the allowable motion are presented in Sec. 5. Tracking experiments are shown in section 6, ranging including complex palm-finger tracking with significant rotation.

2. Hand Model

In our forward dynamics formulation, the hand model (Fig. 1(a)) consists of a base link (palm), and five link-chains (fingers) connected to the base link through five two-degree-of-freedom revolute joints. Each finger is three links connected by two one-degree-of-freedom revolute joints. The finger parts are modeled as cylinders and the palm is modeled as a six-rectangle-side-solid.

A two-degree-of-freedom revolute joint can be simplified as two one-degree-of-freedom revolute joints connected by a zero length and zero mass link, (dummy link)[4]. In the hand model there are 21 links including 5 dummy links and 20 one-degree-of-freedom revolute joints. We number the palm (base link) as link 0. For each finger there are 4 links including one dummy link and 4 joints. The joint connecting the finger to the palm is joint 1, and link 1 connects joint 1 and joint 2 (Link 1 is the dummy link).Joint i connects link $(i-1)$ and link i; link i links joint i and joint $(i+1)$. Each link has a local coordinate frame fixed to its starting end.

The above geometric model is based on the measurements of an average male. The user specifies the joint locations in the image to initialize model finger lengths. When the hand is illuminated by a directional light, we recover surface normal fields of parts of the hand by fitting this basic model (based on previous deformable model methodology that uses shape from shading and edges [19]) to images of the hand. These normals will be used to calculate the generalized shading flow constraint.

3. Image Based Cues

3.1. Fitting the 3-D Model to 2-D images

This approach needs a geometric 3-D model to transform 2-D forces into 3-D ones which will be applied on the dynamic model. Initially the model is fitted to a known pose of the hand, as can be seen in Figure 1(b). At this stage of the work, we assume knowledge of the camera parameters.

At each frame visibility checking is performed in order to match correctly image and model points. The computation of the relative motion to the palm of occluded fingers, is based on the rigid motion of the hand. When the relative motion is not too large, we pick up the finger edges when they reappear. This method will fail when the fingers undergo significant relative motions when occluded. In order to track them successfully in that case, other methods should be integrated in the framework, such as appearance based methods, which is outside the scope of this paper.

3.2. Force Calculation for Dynamic Model

The 3-D finger motion is recovered by fitting the model to image-derived data. The external forces are applied on the dynamic model, then the rotation and translation of finger joints are calculated. Figure 1(c) shows two kinds of typical finger motion. We obtain the forces by calculating displacements using the following procedure.

- Extract the finger edges using the Canny edge operator.
- A curvature-finding operator [7] is used to find the base points of each finger such as B_i, B_j shown in Figure 1(d). The edges between B_i and B_j correspond to the finger segment. The edge points of sub-segments can be derived from the corresponding 3-D points in the 3-D model during tracking.

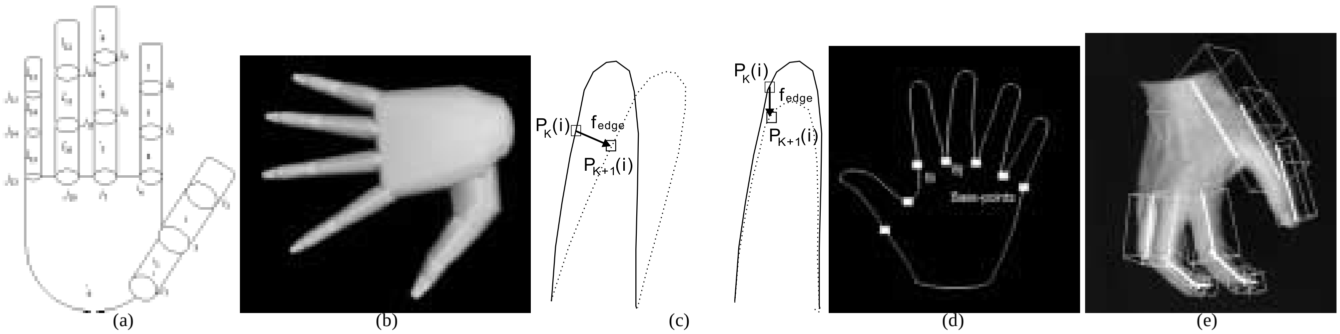


Figure 1. (a) Dynamic Model of Hand. (b) Initial posture of hand model. (c)Finger motion and force from edge displacement. (d)finger segmentation and base points. (e) Representing the projection of the model's articulated segments by their medial axis (thick white line)

- Because the hand motion will result to the change of base-point position between the current- and after-frame, a normalization process is necessary to match the base-points in current- and after- frame according to the distance of two base-points and the length of finger segment.
- Let $p_k(i)$ and $p_{k+1}(i)$ corresponding edge points in k-th frame and $k+1$-th frame. The 2-D force f_{edge} from edge displacement can be calculated by the equation.

$$f_{edge}(i) = p_{k+1}(i) - p_k(i) \tag{1}$$

Another force f_{opt} can be directly derived from the optical flow of the image. In the optical flow equation:

$$I_x u + I_y v + f_t = 0, \tag{2}$$

the temporal differential $\mathbf{e} = (u, v)$ at position (x, y) will be considered as the external force. The optical flow of hand motion is computed by the Lucas-Kanade method[9].

Optical flow near finger edges is not as reliable due to possible mismatches of edge points, so we will only consider the optical flow of the inside area of the finger segment (obtained from the projection of the 3-D model in the image plane). For optical flow computation, we select points with significant gradient magnitude only. In Fig. 2 we see the edge forces and the optical flow forces, applied to different regions of the image.

3.3. Force transformation from 2-D to 3-D

We assume a perspective projection model. Therefore, the point $\mathbf{x} = (x, y, z)$ in the world coordinate system and the point $\mathbf{x_c} = (x_c, y_c, z_c)^T$ in the camera coordinate system ensure the following equation.

$$\mathbf{x} = \mathbf{R_c}\mathbf{x_c} + \mathbf{T_c}, \tag{3}$$

where, $\mathbf{T_c}$ and $\mathbf{R_c}$ are translation and rotation matrices.

Following previous work [8], by taking the time derivatives of the perspective projection equation, with an image point $\mathbf{x_p}$ we get $\dot{\mathbf{x}}p = \mathbf{H}\dot{\mathbf{x}}c = \mathbf{H}(\mathbf{R}_c^{-1}\dot{\mathbf{x}})$, with

$$\mathbf{H} = \left| \begin{array}{ccc} f/z_c & 0 & -x_c/z_c^2 f \\ 0 & f/z_c & -y_c/z_c^2 f \end{array} \right| \tag{4}$$

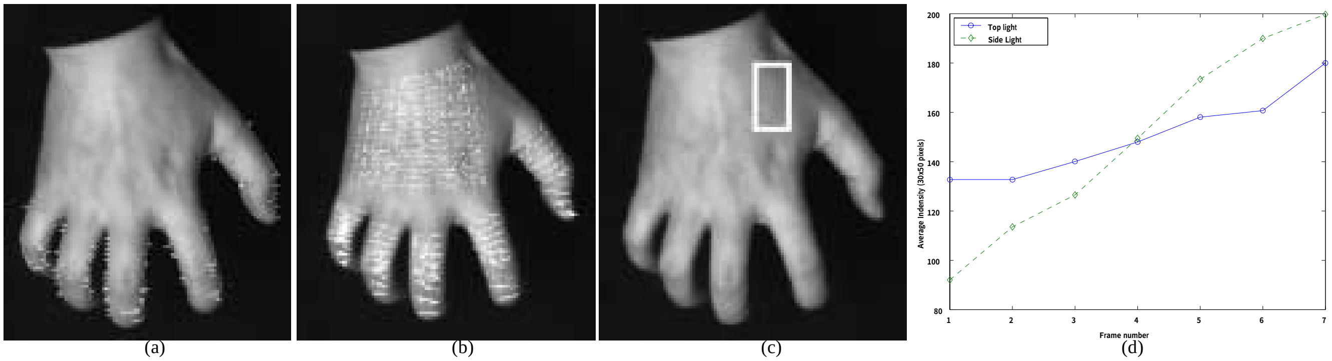
The focal length f is obtained by pre-calibration of the camera. According to deformable model theory these 3D forces are converted to generalized forces $\mathbf{f}_q = \mathbf{J}^\top \mathbf{f}_{3d}$ on the model parameters $\mathbf{q}$, with $\mathbf{J} = \partial \mathbf{x}(x, y, z)/\partial \mathbf{q}$ the Jacobian of the model points, by $\dot{\mathbf{q}} = \mathbf{f}_q$. Consequently, the generalized forces calculated from 2-D images will be $\mathbf{f}_q = (\mathbf{J}_p \mathbf{J})^\top \mathbf{f}_{2d}$ with $\mathbf{J}_p = \mathbf{H}\mathbf{R}_c^{-1}$ the Jacobian of the model points under perspective projection.

To apply the external forces on the dynamic model, we transform the individual forces obtained from edges and the optical flow within every hand segment into one total force and torque to be used in the recursive dynamic framework. The total force and torque for each hand segment are $\mathbf{F} = \sum_{i=1}^{n} f_i, \sum_{i=1}^{n} \mathbf{r_i} \times \mathbf{f_i}$, respectively. $\mathbf{f_i}$ and $\mathbf{r_i}$ are the individual force vector and force position vectors.

4 A new constraint

In previous work [19] a methodology was developed for the incorporation of illumination constraints (any type that is differentiable w.r.t. the model parameters) in a deformable model formulation. In that work, the fitting of the model was done based on a static image, i.e. that data did not change during the fitting process. Hence, any partial derivatives with respect to time in the illumination constraint $\mathbf{C}$ were zero. Here we will generalize our constraint formulation to include image motion. Instead of one image, the fitting process will be guided by a sequence of moving images.

We will start by taking the reflectance equation. Let us assume that we have a reflectance function of the general



(a) (b) (c) (d)

Figure 2. Forces applied to the hand model, and the effects of shading. (a) Edge forces (b) optical flow forces in the interior of the model. (d) is the change in average intensity in a small smooth area of the hand (depicted in (c), when the illumination comes from the top (blue line) and from the side (green dashed line) respectively.

form $I_L = L(\mathbf{l}_p, \mathbf{q})$, where I_L is the observed image intensity and $\mathbf{l}_p$ are the lighting model parameters, which can be differentiated with respect to the normal n of the surface and $\mathbf{q}$ are the hand model parameters. This means that the reflectance of the surface is locally computable and that there are no global illumination effects. We also assume that the illumination parameters do not change with time. The constraint equation is $C = I_L - L(\mathbf{l}_p, \mathbf{q})$, and we differentiate it w.r.t. time, and apply Baumgarte stabilization [3] in order to obtain

$$\dot{\mathbf{C}}(\mathbf{q}, t) + \alpha\mathbf{C} = \mathbf{C}_{\mathbf{q}}\dot{\mathbf{q}} + \mathbf{C}_t + \alpha\mathbf{C} = 0, \tag{5}$$

In this case we cannot ignore the partial derivatives $\mathbf{C}_t$ w.r.t. time. Therefore, using the above formulas we expand Equation 5 to:

$$\frac{\partial \mathbf{I_L}}{\partial \mathbf{q}}\dot{\mathbf{q}} - \frac{\partial \mathbf{L}(\mathbf{l_p}, \mathbf{q})}{\partial \mathbf{q}}\dot{\mathbf{q}} + \frac{\partial \mathbf{I_L}}{\partial t} - \frac{\partial \mathbf{L}(\mathbf{l_p}, \mathbf{q})}{\partial t} + a(I_L - L(\mathbf{l}_p, \mathbf{q})) = 0 \tag{6}$$

We notice that if $\mathbf{J}$ is the Jacobian of the model points, and $\mathbf{J}_p$ is the Jacobian of the model points under perspective projection, as described in Sec. 3, then

$$\frac{\partial \mathbf{I_L}}{\partial \mathbf{q}}\dot{\mathbf{q}} + \frac{\partial \mathbf{I_L}}{\partial t} = \nabla \mathbf{I_L} \mathbf{J}_p\ \mathbf{J}\dot{\mathbf{q}} + \frac{\partial \mathbf{I_L}}{\partial t} \tag{7}$$

is the left hand side of the model based optical flow constraint equation [20]. In model based optical flow, motion field vectors are vectors of velocities of model points, and hence $\dot{\mathbf{x}} = \mathbf{J}\dot{\mathbf{q}}$ applies. Typically in the literature [11] this optical flow term is set to 0. This is correct in the case of ambient only illumination. For the case of light sources at infinity it is also correct for pure translational motion. For the simplest case of a Lambertian surface with a light source at infinity it can be shown [12] that if ω is the angular velocity of the rotational motion and $\mathbf{l}$ the light source direction, the magnitude of the error between the true motion field and the apparent (and computable) optical flow is

$$|Dv| = \rho \frac{|\mathbf{l}(\omega \times \mathbf{n})|}{||\nabla E||} \tag{8}$$

This error is small when the change of gradient is big, but in the case of smooth surfaces this effect becomes much more pronounced. Similarly $\frac{\partial \mathbf{L}(\mathbf{l_p}, \mathbf{q})}{\partial t} = 0$ since normals change based only on the model parameters $\mathbf{q}$.

This means that when there is no motion the constraint equation simplifies to the shading constraint. Therefore

$$\frac{\partial \mathbf{I_L}}{\partial \mathbf{q}}\dot{\mathbf{q}} - \frac{\partial \mathbf{L}(\mathbf{l_p}, \mathbf{q})}{\partial \mathbf{q}}\dot{\mathbf{q}} + \frac{\partial \mathbf{I_L}}{\partial t} - a(L(\mathbf{l}_p, \mathbf{q}) - I_L) = 0 \tag{9}$$

encompasses both constraints. In the case of a smooth moving object (9) allows to deal with errors due to directed illumination and offers the possibility of recovering the motion of relatively smoothly shaded surfaces. Fig. 2(c), (d) shows the change in average intensity in a small smooth area of the hand, when the illumination comes from the top and from the side respectively. In the second case, changes in the intensity of the points are dramatic.

5 Dynamic Tracking of Hand Motion

In our methodology we estimate the hand motion in response to the applied 3D forces on the hand as a Forward Dynamics problem where given the external forces we want to compute the velocity and position of the palm.

Since we use a recursive dynamic formulation we will use Featherstone's[2], [5] spatial notation to model our kinematic and dynamic variables. We integrate the constraint of Eq. 9 in the above formulation to determine the vector $\mathbf{q}$ of the model's degrees of freedom which includes the joint variables, global rotation and translation.

Furthermore, human fingers are not ideal dynamic links, their joints have upper and lower bounds. Therefore, we need to solve the above dynamic equations under joint limit constraints. These joint limits which constrain the relative motion of fingers together with our dynamic formulation which does not allow the inter-penetration of fingers make hand tracking significantly more robust. Our method has the following steps:

1. At time t, mark the joints that reach their joint limits.
2. Solve the dynamic equations of the hand at time $t + dt$ recursively.
3. For each finger, starting at joint 1 (the joint that connects the palm and the finger), mark the first joint that keeps at its joint limit during the time period from t to $t + dt$. If there is no such joint, go to step 6.
4. Fix the joints marked at step 3, and merge two links connected by a fixed joint to one link. Update the dynamic hand model.
5. Go back to step 2.
6. Output the status of the dynamic model of the hand at time $t + dt$. Increase time $t = t + dt$, and go to step 1.

6. Experiments

We performed a series of experiments to test our method with a variety of hand motions. All our experiments run on a PIII 500MHz processor at approximately 4 frames per second. Two similar datasets were taken under two different illumination conditions. The first dataset (Fig. 3) was taken with the light coming from the top of the hand, thus minimizing the variations in intensity w.r.t. the illumination. The second dataset (Fig. 4) was taken with the light on the side (approx 50 degrees) so illumination effects are pronounced. Each sequence was approximately 100 frames. Due to space limitations we include only a few frames in this paper. The full sequences and the tracking results are available as movie files `http://www.cs.sunysb.edu/~samaras/hand/`. Files `trk_top.mpg` and `trk_side.mpg` respectively. To show the accuracy of the tracking we project the segments of the hand model back onto the image. We represent the segments by their medial axes (Fig. 1 (e)). At the same web site, an additional data sequence where fingers flex to a closed position and unflex back to open without losing track is in movie file `trk_flex.mpg` and the full model while tracking but rendered from a different viewpoint in movie file `mdl_flex.mpg`. From such a viewpoint, it can be seen that our dynamic model allows for accurate tracking of segments that are almost occluded from the camera.

In figure 3, we present a number of complex rotational motions for the fingers and for the whole hand. First the fingers bend away from the camera, then the whole hand rotates with significant occlusions. Neither edges nor optical flow alone would have succeeded in tracking this sequence. Finally in figure 4 we demonstrate the increased power of the shading flow constraint, since classic optical flow based on the brightness constancy assumption fails due to the significant appearance changes from frame to frame due to illumination. We notice that tracking is quite successful in these examples. There are some slight inaccuracies tracking the thumb, here modeled as a 3 segment finger with one segment of zero lebgth, whereas a better model would have 2 segments only

7. Conclusions

In this paper we have augmented traditional optical flow and replaced it with a more general equation that includes shading information. We have used this formulation within a deformable model framework and we were able to track difficult hand motions under a variety of illumination conditions. To improve the efficiency of the approach we use the augmented equations only in areas where the optical flow constraint is significantly violated. Our dynamic hand model formulation allows the integration of multiple cues and for robustness we also use edges in our tracking. We have shown tracking results for simple and complex palm and finger motions. Future work includes better occlusion recovery handling using Kalman Filtering and the incorporation of other sources of visual information such as color, in order to work on cluttered backgrounds.

References

[1] G. Engeln-Mullges, F. Uhlig. Numerical Algorithms with C. *Springer*, 1996

[2] R. Featherstone. Robot Dynamics Algorithm. *Kluwer Academic*, Boston, 1987

[3] J. Baumgarte. Stabilization of constraints and integrals of motion in dynamical systems. *Computer Methods in Applied Mechanics and Engineering*, 1:1–16, 1972.

[4] G. Huang, D. Metaxas, and J. Lo. Human Motion Planning Based on Recursive Dynamics and Optimal Control Techniques. *Computer Graphics International 2000*, pp. 19-28.

[5] K.W. Lilly. Efficient Dynamic Simulation of Robotic Mechanisms. *Kluwer Academic*, Boston, 1993

[6] J. Lo and D. Metaxas. Recursive Dynamics and Optimal Control Techniques for Human Motion Planning. *CA'99*. Geneva, Switzerland, May 26-29, 1999

[7] S.B. Kang and K. Ikeuchi, "Toward Automatic Instruction from Perception – Recognizing a Grasp from Observation", *IEEE Trans. of Robotics and Automation*, pp.432-443, Aug. 1993.

[8] D.N. Metaxas, *Physics-Based Deformable Models – Applications to Computer Vision, Graphics and Medical Imaging*, Kluwere Academic Publishers, 1998.

[9] B. Lucas and T. Kanade, "An Iterative Technique of Image Registration and Its Application to Stereo",*Proc. 7th IJCAI*, pp.674-679, August, 1981.

[10] J. Angeles and O. Ma. Dynamic Simulation of n-Axis Serial Robotic Manipulators Using a Natural Orthogonal Complement. *The International Journal of Robotics Research*, 7(5):32-47, October 1988.

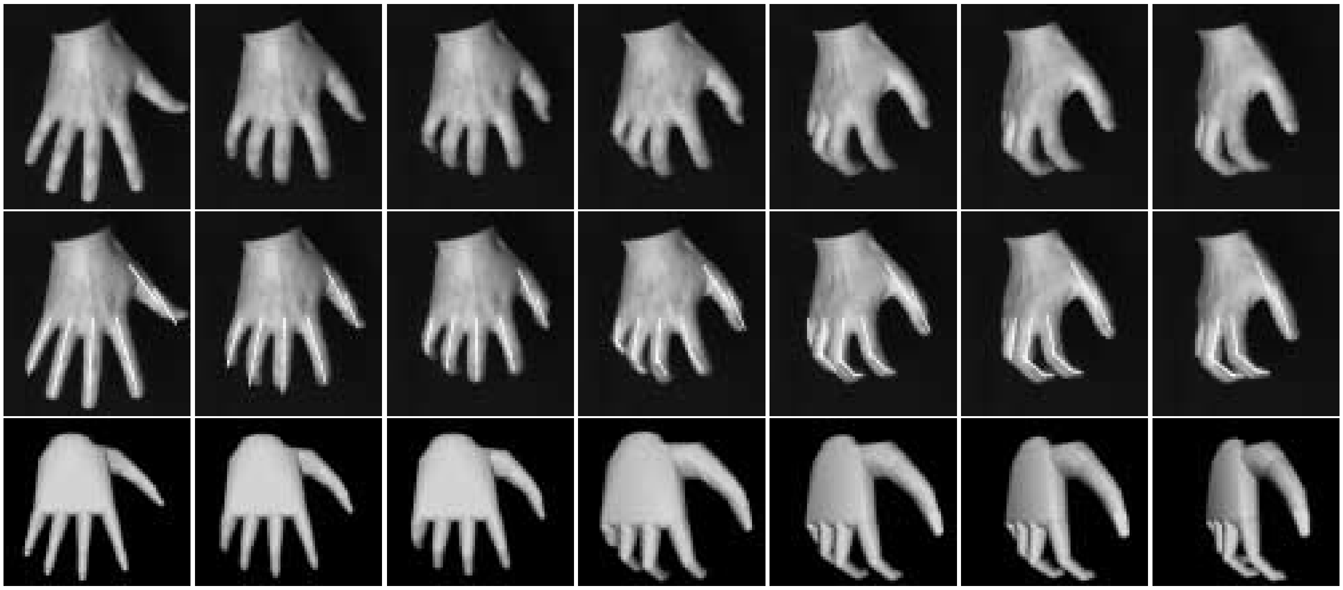


Figure 3. Seven frames from a longer sequence tracking flexing of fingers and hand rotation. First row: Original data. Second row: The accuracy of the tracking is demonstrated by projecting the medial axes of each model finger (white lines) on the tracked data. Third row: Full model during tracking. The full sequence can be seen in movie clip file `http://www.cs.sunysb.edu/~samaras/hand/trk_top.mpg`

[11] B.K.P. Horn. *Robot Vision*. 1986.

[12] A. Verri and T.A. Poggio. Motion field and optical flow: Qualitative properties. *PAMI*, 11(5):490–498, May 1989.

[13] H. Brandl, R.Johanni, and M.Otter. An Algorithm for the Simulation of Multibody Systems with Kinematic Loops. *IFToMM Seventh World Congress on the Theory of Machines and Mechanisms*, Sep. 1987.

[14] J.K. Hodgins, W.L. Wooten, D.C. Brogan, and J.F. O'Brien. Animation of Human Athletics. *SIGGRAPH 95*.

[15] R. H. Lathrop. Constrained (Closed-Loop) Robot Simulation by Local Constaint Propagation. *IEEE ICRA '86*.

[16] A.J. Stewart and J.F. Cremer. Beyond keyframing: An algorithmic approach to animation. *Graphics Interface*, 1992

[17] M. W. Walker and D. E. Orin. Efficient Dynamic Computer Simulation of Robotic Mechanisms. *Journal of Dynamic Systems, Measurement and Control*,104:205-211, Sep. 1982.

[18] J. Wilhelms and B. Barsky. Using Dynamic Analysis to Animate Articulated Bodies such as Humans and Robots. In *Graphics Interface*, 1985

[19] D. Samaras and D. Metaxas. Incorporating Illumination Constraints in Deformable Models. CVPR 1998, pp. 322-329.

[20] D. DeCarlo and D. Metaxas. Optical Flow Constraints on Deformable Models with Applications to Face Tracking. IJCV, July 2000, 38:2, pp. 99-127.

[21] Negahdaripour, S., Revised Definition of Optical Flow: Integration of Radiometric and Geometric Cues for Dynamic Scene Analysis, PAMI(20), No. 9, Sep. 1998, pp. 961-979

[22] Carceroni, R.L., Kutulakos, K.N, Multi-View Scene Capture by Surfel Sampling: From Video Streams to Non-Rigid 3D Motion, Shape and Reflectance, ICCV01(II: 60-67)

[23] Haussecker, H., and D. J. Fleet, 2000. Computing optical flow with physical models of brightness variation, PAMI(23), No. 6, pp. 661-673, 2001.

[24] J J Kuch and T. S. Huang. Vision-based hand modeling and tracking for virtual teleconferencing and telecollaboration. In ICCV95, pg.666-671,1995.

[25] J. Lee and T. Kunii. Model-based analysis of hand posture. IEEE CGA, 15:77-86, Sept. 1995.

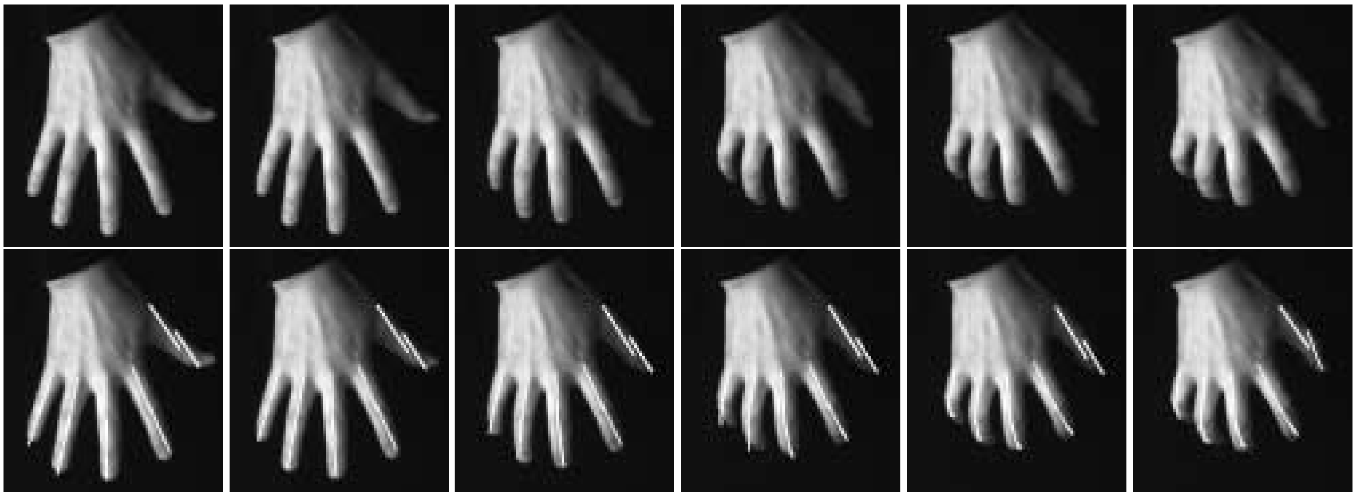
[26] J. Rehg and T. Kanade. Model-based tracking of self-occluding articulated objects. In ICCV '95, pg 612-617.

[27] Ying Wu and T. S. Huang. Capturing articulated human hand motion: A divide-and-conquer approach. In ICCV '99, pg 606-611, Corfu, Greece, Sept. 1999.

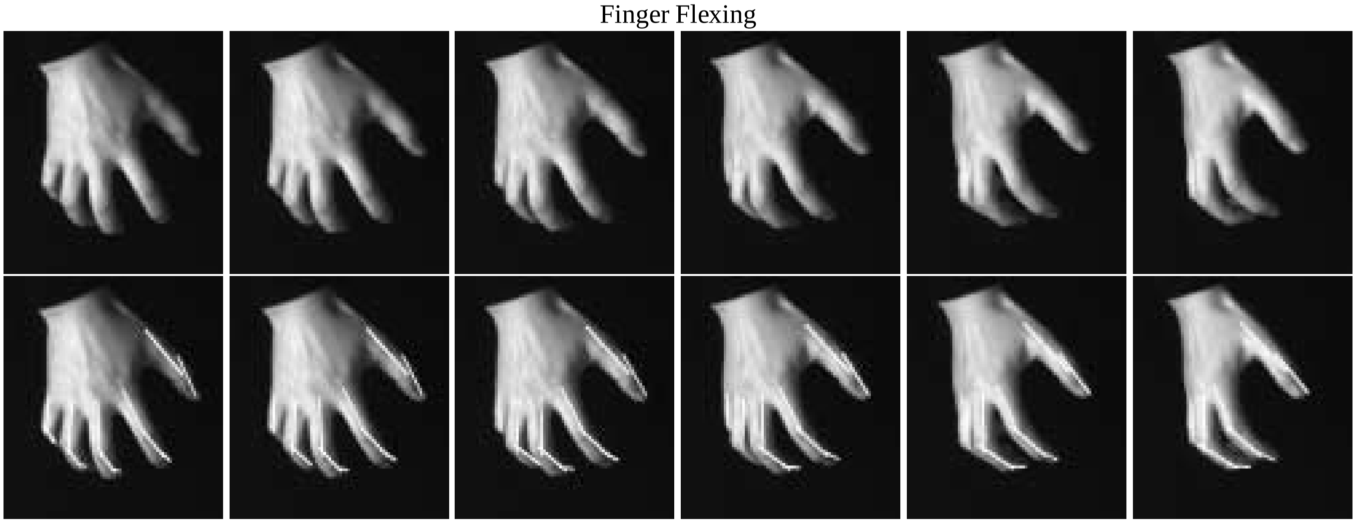
[28] Ying Wu; Lin, J.Y.; Huang, T.S. Capturing natural hand articulation. ICCV '01 (II:426-432)

[29] Rosales, R., Athitsos, V., Sigal, L., and Sclaroff, S., 3D Hand Pose Reconstruction Using Specialized Mappings, ICCV01

[30] Q. Delamarre and O. Faugeras, "Finding pose of hand in video images: a stereo-based approach", AFGR 98.



Finger Flexing



Hand Rotation

Figure 4. Eight frames from a longer sequence tracking flexing of fingers and hand rotation. Sideways illumination causes significant deviations from classical optical flow constraint during rotation. The generalized optical flow constraint with shading allows for accurate tracking. First and third row: Original data. Second and fourth row: The accuracy of the tracking is demonstrated by projecting the medial axes of each model finger (white lines) on the tracked data. The full sequence can be seen in movie clip file `http://www.cs.sunysb.edu/~samaras/hand/trk_side.mpg`**.**