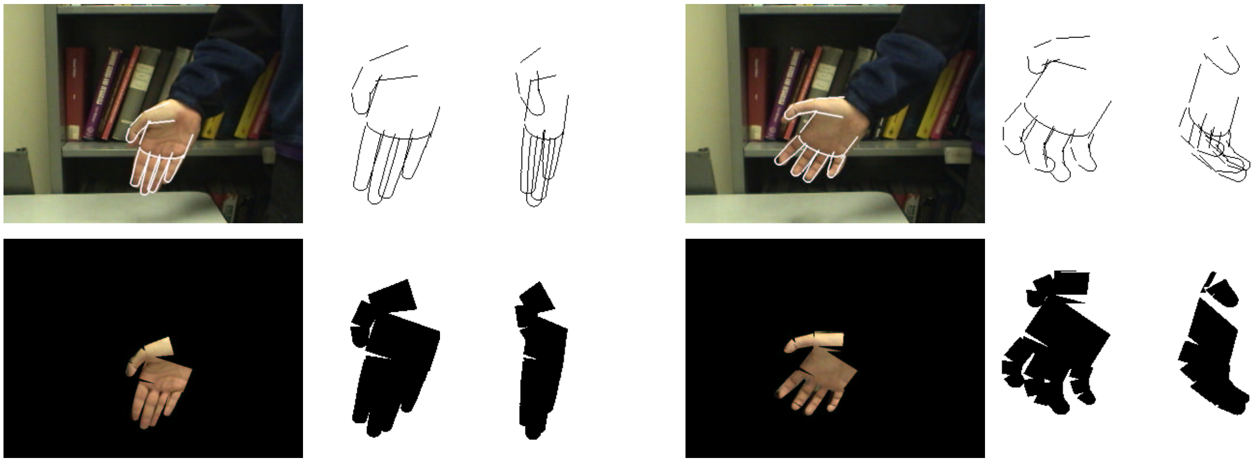

Fig. 1. Projected edges (top row) and silhouettes (bottom row) for two configurations (left and right blocks) of the 3D structural hand model. To aid visualization, the model joint angles are set to match the images (left), and then also projected following rotations by 35° (center) and 70° (right) about the vertical axis.

a graphical model for the tracking problem, we consider a redundant *local* representation in which each hand component is described by its own three–dimensional position and orientation. We show that the model's kinematic constraints, including self–intersection constraints not captured by joint angle representations, take a simple form in this local representation. Furthermore, in cases where model components do not significantly occlude each other, standard edge and color based likelihood measures may be similarly decomposed. We describe the implementation of NBP on this model, as well as several methods for improving computational efficiency. These include a novel method for fast orientation–based Chamfer distance evaluation using KD–trees [2]. We conclude with simulations demonstrating that NBP can refine noisy initializations in single frames, as well as track hand motion over two extended sequences.

II. Geometric Hand Modeling

A. Structural Model

Structurally, the hand is composed of sixteen approximately rigid components: three phalanges or links for each finger and thumb, as well as the palm [25]. As proposed by [14, 19], we model each rigid body by one or more truncated quadrics (ellipsoids, cones, and cylinders). These geometric primitives are well matched to the true geometry of the hand, and in contrast to 2.5–dimensional "cardboard" models [24, 26], allow tracking from arbitrary viewing orientations. In addition, because the perspective projection of a quadric surface is a conic, one can efficiently determine the image points lying on the boundary or silhouette of the projection of any three–dimensional model configuration [3, 19].

Figure 1 shows the edges and silhouettes corresponding to two different configurations of the hand model, each of which is seen from three different viewpoints. Because our model is designed for estimation, not visualization, precise modeling of all parts of the hand is unnecessary. As our tracking results demonstrate, it is sufficient to capture the coarse structural features which are most relevant to the observation model described in Sec. III. Note also that we do not consider model self–occlusion when finding edges. See Sec. III-A for further discussion of this approximation.

B. Kinematic Model

The kinematic constraints between different hand model components are well described by revolute joints [25]. Figure 2(a) shows a graph describing this kinematic structure, in which nodes correspond to rigid bodies and edges to joints. The two joints connecting the phalanges of each finger and thumb have a single rotational degree of freedom, while the joints connecting the base of each finger to the palm have two degrees of freedom (corresponding to grasping and spreading motions). Thus, twenty joint angles are required to describe the relative positions of all hand parts.

The full configuration of the hand is described by these angles along with the palm's global position and orientation, giving a total of 26 degrees of freedom. Given image measurements, calculation of a model configuration's likelihood generally requires the global position and orientation of each component. This forward kinematics problem is easily solved via a series of transformations derived from the position and orientation of each joint axis, along with the corresponding joint angles (see, for example, [12] for details).

C. Redundant Local State Representation

Most model–based hand trackers parameterize the model state in terms of the twenty joint angles described above, along with the palm's global position and orientation. In this paper, we instead explore a redundant representation in which the i^{th} rigid body is described by its position q_i and orientation r_i (a unit quaternion). Let $x_i = (q_i, r_i)$ denote this *local* description of each hand component's configuration, and $x = \{x_1, \ldots, x_{16}\}$ the configuration of the entire hand.

Clearly, there are dependencies among the elements of x implied by the kinematic constraints. Let $\mathcal{E}_K$ be the set of all