

The Institute For Research In Cognitive Science

Active Motion-Based Segmentation of Human Body Outlines

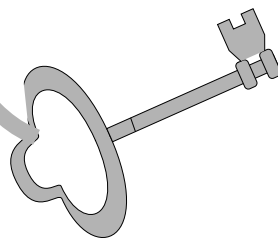
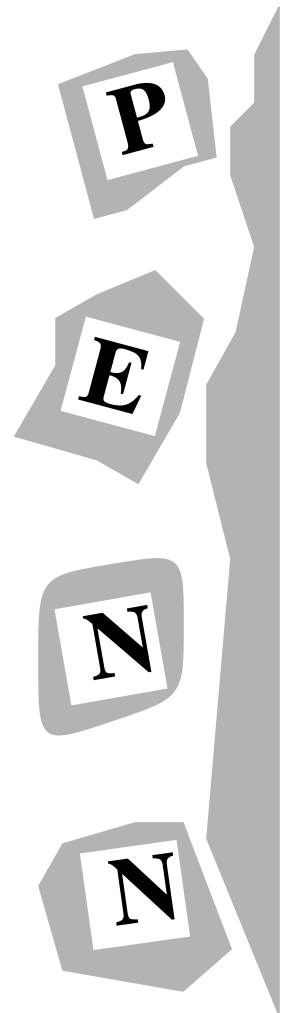
by

**Ioannis A. Kakadiaris
Dimitri Metaxas
Ruzena Bajcsy**

**University of Pennsylvania
3401 Walnut Street, Suite 400C
Philadelphia, PA 19104-6228**

November 1994

Site of the NSF Science and Technology Center for
Research in Cognitive Science



Active Motion-Based Segmentation of Human Body Outlines

Ioannis A. Kakadiaris, Dimitri Metaxas and Ruzena Bajcsy

GRASP Laboratory

Department of Computer and Information Science

University of Pennsylvania

Philadelphia, PA 19104

Abstract

We present an integrated approach towards the segmentation and shape estimation of human body outlines. Initially, we assume that the human body consists of a single part, and we fit a deformable model to the given data using our physics-based shape and motion estimation framework. As an actor attains different postures, new protrusions emerge on the outline. We model these changes in the shape using a new representation scheme consisting of a parametric composition of deformable models. This representation allows us to identify the underlying human parts that gradually become visible, by monitoring the evolution of shape and motion parameters of the composed models. Based on these parameters, their joint locations are identified. Our algorithm is applied iteratively over subsequent frames until all moving parts are identified. We demonstrate our technique in a series of experiments with very encouraging results.

1 Introduction

The task of human motion analysis can be decomposed into the following subtasks: (1) estimation of the shape and motion parameters of the parts of a human body, and (2) recognition of the motion performed by the actor. The estimation subtask is important in applications such as anthropometry, human factors design, ergonomics, performance measurement of both athletes and patients with psychomotor disabilities and virtual reality. On the other hand, the recognition subtask is important in applications such as human computer interaction.

The human body consists of non-rigid articulated parts whose motion depends on joints with various shapes and on intricate muscle actions. To recover the degrees of freedom associated with the shape and motion of a moving human body, many researchers either use a model-based approach [10, 4, 1, 7, 8, 15, 14] or employ certain assumptions [13, 18, 3, 17, 2, 19, 6]. Most of these techniques assume models that can only approximate the human body (e.g., generalized cylinders) and cannot adapt to different body sizes, since they are not deformable. To overcome this

problem other researchers assume prior segmentation of the given data into parts [12] and then fit deformable models that can adapt to data from humans of different sizes [11, 9]. However, the process of segmentation and the process of shape and motion estimation are decoupled leading to possible lack of robustness and inaccuracies. Moreover, no techniques exist that acquire a concise model of the human body automatically. The solution to the above problems requires the development of an algorithm that: (1) integrates the processes of segmentation and fitting, (2) allows reliable shape description of the parts, (3) estimates the true location of the joints between the parts, (4) detects multiple joints, and (5) obviates the need for markers and special equipment.

As a first step towards such an algorithm, we have presented a physics-based approach to the shape and motion estimation of non-occluded chain-like structures in human body outlines (e.g., arms and legs) [5]. In this paper, we extend our technique to be able to fully segment outlines of moving humans. Initially, we assume that the human body consists of a single part. Using our physics-based framework, we fit a deformable model to the given time-varying data and we monitor the relevant model parameters. Due to intra-part occlusion (partial or full) and the relative motion of the human parts, the shape of the outline dynamically changes. In particular, tree-like structures (e.g., from the motion of the arm with respect to the torso) and chain-like structures (e.g., from the motion of the upper arm with respect to the forearm) are protruding from the outline. Identifying these structures will lead to the identification of the underlying human parts. To this end, we introduce a new method for modeling shapes with large protrusions which amounts to composition of deformable models. This method allows us to represent the shape of an outline in a compact way and to hypothesize an underlying part structure. We can verify the hypothesis by monitoring the relative motion of the composed deformable models. Once the hypothesis of multiple underlying parts is verified, we are able to identify their joint location based on the estimated shape and motion parameters. Our algorithm

for part-identification, and for the estimation of shape and motion is applied iteratively over subsequent frames until all the moving parts are identified.

In this paper, we first formulate the theory of parametric composition of geometric primitives. We then present the algorithm for shape and motion estimation for parts that gradually become visible. Finally, we present selected experimental results demonstrating our algorithm.

2 Deformable models

Following the notation in [9], the 3D position \mathbf{x} of a point (with material coordinates $\mathbf{u} = (u, v)$) on a deformable model at time t , with respect to an inertial frame of reference Φ is given by $\mathbf{x}(\mathbf{u}, t) = \mathbf{c}(t) + \mathbf{R}(t)\mathbf{p}(\mathbf{u}, t)$, where $\mathbf{p}(\mathbf{u}, t)$ is the position of the given point with respect to a model-centered reference frame ϕ . In addition, $\mathbf{c}(t)$ and $\mathbf{R}(t)$ denote the position of the origin and the orientation of ϕ with respect to Φ . We further express the position \mathbf{p} as: $\mathbf{p} = \mathbf{s} + \mathbf{d}$, where $\mathbf{s}(\mathbf{u}, t)$ is the model's global reference shape and $\mathbf{d}(\mathbf{u}, t)$ represents a displacement function. We define the global reference shape as $\mathbf{s} = \mathbf{T}(\mathbf{e}(\mathbf{u}; a_0, a_1, \dots); b_0, b_1, \dots)$, where the geometric primitive \mathbf{e} (defined parametrically over \mathbf{u} with global shape parameters a_i) is subjected to the global deformation \mathbf{T} which depends on the parameters b_i . To represent the local deformations, we employ finite element shape functions which are tensor products of one-dimensional Hermite polynomials [5]. Therefore, $\mathbf{d} = \mathbf{S}\mathbf{q}_d$, where \mathbf{S} is the shape matrix whose entries are the shape functions and $\mathbf{q}_d = (\dots, d_i^\top, \dots)^\top$ is the vector of local deformation parameters.

2.1 Parametric Composition of Geometric Primitives

To represent large protrusions or concavities and their shape evolution in a compact and intuitive way, we introduce a new representation based on the parametric composition of primitives. Intuitively, using the parametric composition of primitives, we can describe compactly the shape of their union (in the case of protrusions) or intersection (in the case of concavities). Fig. 1 depicts an example of composition (union) of two superellipsoids.

In the interest of space, we will formulate the theory of composition in 2D, to represent the shape of the boundary of the union of primitives¹. Let \mathbf{x}_0 and \mathbf{x}_1 be two 2D parametric primitives (defined by the mappings $\mathcal{C}_0: [v_{0_b}, v_{0_e}] \rightarrow \mathbb{R}^2$ and $\mathcal{C}_1: [v_{1_b}, v_{1_e}] \rightarrow \mathbb{R}^2$ respectively - where the subscripts b and e denote the beginning and end of the domain) positioned in space so that \mathbf{x}_1 , *intersecting primitive*, intersects \mathbf{x}_0 , *root primitive*, at points

¹In the 2D case the material coordinate space is one-dimensional and $\mathbf{u} = v$.

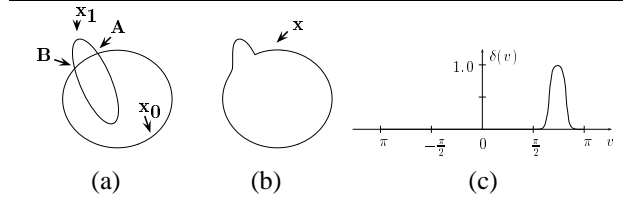


Figure 1: An example of composition of two superellipsoids: (a) depicts $\mathbf{x}_1(v_1)$ which intersects $\mathbf{x}_0(v_0)$ at points A and B, (b) depicts their composition $\mathbf{x}(v)$, and (c) depicts the composition function $\delta(v; 0.65\pi, 0.77\pi, 10)$.

A and B. The coordinates of these two points can be expressed in terms of either the material coordinate v_0 of \mathbf{x}_0 or the material coordinate v_1 of \mathbf{x}_1 . Let $v_{0,A}$ and $v_{0,B}$ be the values of v_0 , and $v_{1,A}$ and $v_{1,B}$ be the values of v_1 at the points A and B, respectively (Fig. 2(a,d)). Without loss of generality, we can assume that we name the points of intersection A and B so that the relation $v_{0,A} < v_{0,B}$ holds. Also, let I_0 be the curve segment of \mathbf{x}_0 which lies in the interior of the union of \mathbf{x}_0 and \mathbf{x}_1 (Fig. 2(b)), and J_0 be the curve segment of \mathbf{x}_0 which belongs to the boundary of their union (Fig. 2(c)). We define I_1 and J_1 in a similar way (Fig. 2(e,f)). Intuitively, in composing two primitives to represent the boundary of their union, we want to map I_0 to J_1 and J_0 to I_1 . However, depending on the position of the point $\mathcal{C}_0(v_{0_b})$ (if it belongs or not to I_0), the curve I_0 can be the map of either a continuous interval or of a union of continuous intervals. For example, a superellipsoid $\mathbf{x}_0(v_0)$ is defined by the mapping $\mathcal{C}_0: [-\pi, \pi] \rightarrow \mathbb{R}^2$ as depicted in Fig. 2(c). When a superellipsoid \mathbf{x}_1 intersects \mathbf{x}_0 , the point $\mathcal{C}_0(-\pi)$ either belongs to I_0 (Fig. 2(i,k)) or not (Fig. 2(h,j)). In the first case, I_0 is the map of the union of two continuous intervals $I_0 = \{\mathbf{x}_0(v_0) : v_0 \in (v_{0,B}, \pi) \cup [-\pi, v_{0,A}]\}$. In the second case, I_0 is the map of a single interval $I_0 = \{\mathbf{x}_0(v_0) : v_0 \in (v_{0,A}, v_{0,B})\}$. This distinction arises from the fact that the open interval $[v_{0_b}, v_{0_e})$ and the closed curve \mathbf{x}_0 are not homeomorphic².

Based on the above, the shape \mathbf{x} of the composed primitive ($\mathcal{C}: [v_b, v_e] \rightarrow \mathbb{R}^2$), can be defined in terms of the parameters of the defining primitives \mathbf{x}_0 and \mathbf{x}_1 as follows:

$$\mathbf{x}(v) = (1 - \delta(v)) \mathbf{x}_0(h_0(v)) + \delta(v) \mathbf{x}_1(h_1(h_0(v))), \quad (1)$$

where $\delta: [v_b, v_e] \rightarrow [0, 1]$ is the composition function for the union of two primitives. Specifically, $\delta(v; v_{\min}, v_{\max}, c) = \sigma(c(v - v_{\min})) - \sigma(c(v - v_{\max}))$, where c is a constant that controls the shape of the function δ at the neighborhoods of v_{\min} and v_{\max} , and $\sigma: \mathbb{R} \rightarrow [0, 1]$ is defined as: $\sigma(x) = \frac{1 + \tanh(x)}{2}$. The function σ approximates the step function. The piecewise linear function $h_0: [v_b, v_e] \rightarrow [v_{0_b}, v_{0_e}]$ maps the material coordinate v to v_0 and is defined as: $h_0(x) = f_1(x)$ (see appendix for the

²A set \mathbf{S} is topologically equivalent or homeomorphic to a set \mathbf{T} iff there is a 1-1 bicontinuous mapping f of \mathbf{S} onto \mathbf{T} .

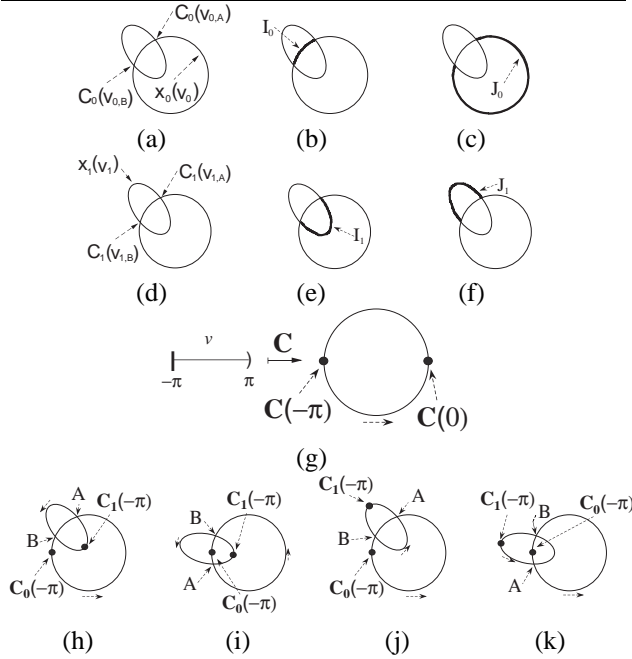


Figure 2: Figs.(a-c) illustrate the notation pertaining to \mathbf{x}_0 , Figs.(d-f) illustrate the notation pertaining to \mathbf{x}_1 , Fig.(g) demonstrates that the interval $[-\pi, \pi]$ and a closed curve are not topologically equivalent, and Figs.(h-k) depict possible positions of the points $C_0(-\pi)$ and $C_1(-\pi)$.

definition of the functions $f_i(x), i = 1 \dots 4$. The piecewise linear function $h_1(x): [v_{0_b}, v_{0_e}] \rightarrow [v_{1_b}, v_{1_e}]$ maps the material coordinate v_0 to v_1 . In effect, $h_1(x)$ maps I_0 to J_1 and J_0 to I_1 in order to form the union. For the definition of $h_1(x)$, four cases have to be distinguished. These four cases reflect if the curve segments I_0, J_0, I_1 and J_1 (their definition for each case are given at the appendix) are maps of a single continuous interval or of a union of two continuous intervals.

Case I: If $C_0(v_{0_b}) \in J_0$ and $C_1(v_{1_b}) \in I_1$ (Fig. 2(h)), then

$$h_1(x) = \begin{cases} f_1(x) & x \in (v_{0,A}, v_{0,B}) \\ f_4(x) & x \in [v_{0,B}, \pi] \cup [-\pi, v_{0,A}]. \end{cases}$$

The function $f_1: (v_{0,A}, v_{0,B}) \rightarrow (v_{1,A}, v_{1,B})$ maps I_0 to J_1 , and the function $f_4: [v_{0,B}, \pi] \cup [-\pi, v_{0,A}] \rightarrow [v_{1,B}, \pi] \cup [-\pi, v_{1,A}]$ maps J_0 to I_1 .

Case II: If $C_0(v_{0_b}) \in I_0$ and $C_1(v_{1_b}) \in I_1$ (Fig. 2(i)), then

$$h_1(x) = \begin{cases} f_2(x) & x \in (v_{0,B}, \pi) \cup [-\pi, v_{0,A}] \\ f_3(x) & x \in [v_{0,A}, v_{0,B}]. \end{cases}$$

The function $f_2: (v_{0,B}, \pi) \cup [-\pi, v_{0,A}] \rightarrow (v_{1,B}, v_{1,A})$ maps I_0 to J_1 , and the function $f_3: [v_{0,A}, v_{0,B}] \rightarrow [v_{1,A}, \pi] \cup [-\pi, v_{1,B}]$ maps J_0 to I_1 .

Case III: If $C_0(v_{0_b}) \in J_0$ and $C_1(v_{1_b}) \in J_1$ (Fig. 2(j)), then

$$h_1(x) = \begin{cases} f_3(x) & x \in (v_{0,A}, v_{0,B}) \\ f_2(x) & x \in [v_{0,B}, \pi] \cup [-\pi, v_{0,A}]. \end{cases}$$

The function $f_3: (v_{0,A}, v_{0,B}) \rightarrow (v_{1,A}, \pi) \cup [-\pi, v_{1,B})$ maps I_0 to J_1 , and the function $f_2: [v_{0,B}, \pi] \cup [-\pi, v_{0,A}] \rightarrow [v_{1,B}, v_{1,A}]$ maps J_0 to I_1 .

Case IV: If $C_0(v_{0_b}) \in I_0$ and $C_1(v_{1_b}) \in J_1$ (Fig. 2(k)), then

$$h_1(x) = \begin{cases} f_4(x) & x \in (v_{0,B}, \pi) \cup [-\pi, v_{0,A}] \\ f_1(x) & x \in [v_{0,A}, v_{0,B}]. \end{cases}$$

The function $f_4: (v_{0,B}, \pi) \cup [-\pi, v_{0,A}] \rightarrow (v_{1,B}, \pi) \cup [-\pi, v_{1,A})$ maps I_0 to J_1 , and the function $f_1: [v_{0,A}, v_{0,B}] \rightarrow [v_{1,A}, v_{1,B}]$ maps J_0 to I_1 .

The equations above generalize easily to the case of multiple intersecting primitives. In the following section, we will examine how the proposed representation can be used for the analysis of human outlines.

3 Active body part extraction: shape and motion estimation

The goal of our approach is to automatically segment body outlines of moving humans without using a prior model of the human body and the shape of its parts. To accomplish this goal, we request that the actor performs a set of motions according to a protocol that reveals the structure of the human body. All the motions start and end in the position where the body is erect and the arms are placed straight against the sides of the body facing medially (*body reference position*). The set of motions required to identify the shape and connectivity of the head and the limbs of an actor are described below.

Protocol of Motions

- 1. Head Motion:** The actor tilts the head forward as far as possible.
- 2. Left (or right) upper body extremities motions:** In the first phase of the motion, the actor lifts the left (or right) arm to the front until the arm reaches the horizontal position. Continuing from this position to the second phase, the actor rotates the arm so that the hand is facing downwards and flexes the wrist. Then, the actor bends the elbow, bringing the forearm to the vertical position.
- 3. Right (or left) upper body extremities motions:** The actor lifts the left (or right) arm backwards to a comfortable position, in which the arm is not fully occluded by the torso. Then, the actor performs the left upper body extremities motions using the right arm.
- 4. Left (or right) lower body extremities motions:** This motion consists of two phases. In the first phase, the actor extends the left (or right) leg to the front. When the leg reaches the maximum comfortable position (in which the