Medial Axis Representations for the Analysis of Human Shape in Images

Charlie Kemp

The Problem: Since the 1960's researchers have observed that biological shape can be usefully described in terms of medial axes [1]. Unfortunately, the traditional algorithms used to estimate medial axis descriptions of shape are of limited use due to a variety of practical problems, including instability to minor perturbations of the input [2].

Motivation: By using more recent signal processing techniques in conjunction with an updated representational framework, the benefits of medial axis representations can be exploited while avoiding several of their traditional drawbacks. Some types of medial axis representations appear to be well matched to the class of shapes that comprise human faces and bodies. These descriptors highlight the underlying structure of the overall shape, while also representing important multi-scale aspects of the component shapes. As an intuitive example, a stick figure representation of a human body does a good job of encoding the body's configuration with respect to the world and the observer. Furthermore, each part of the stick figure can describe a component shape of the body at some scale, such as an elongated rectangular shape for the torso and legs, and two smaller tapering tubes for the legs.

Previous Work: Several groups have demonstrated that signal processing techniques can be used to construct stable medial axis representations [3] [7]. 3D skeletal representations, a 3D generalization of the 2D medial axis, have been shown to be well suited to modeling the surface structure and kinematics of flora and fauna [6] [5]. This property of many shapes in nature derives directly from the common mechanisms of morphological development in biology. For example, limb buds, growth points, are often used in the development of animal appendages and plant branches. Essentially, each limb bud grows along a 3D curve that approximately forms a 3D medial axis for the animal's appendage or the plant's branch [1] [9]. Often the projections of these 3D axes onto an image plane are nearly identical to the 2D axes derived from the 3D shape's projection onto the image plane [10]. Consequently, the existence of useful 3D skeletal representations in biological objects often leads to useful 2D skeletal representations over a large set of viewing angles. Since medial axes are parts-based representations they have an accompanying structural representation that indicates how the component parts are put together. This structural level of representation is relatively invariant to changes in the viewpoint of the camera as well as the kinematic state of complex articulated shapes, such as the human body [8].

Approach: First, we have created new analytical descriptions of the medial axis features we are attempting to estimate. Second, we are working on signal processing mechanisms that facilitate the robust estimation of these features.

Difficulty: Many applications require real-time or near real-time performance. Without significant attention to approximations, optimizations and specialization of the general algorithms with which we are working, these new methods will not perform in real-time using a standard general-purpose workstation.

Impact: Improving and adding to the theory and algorithms available for the robust estimation of 2D medial axes from real images will help produce useful methods for analyzing biological shape in images. These methods should be particularly good at detecting and interpreting human shape in applications involving human-machine interaction.

Future Work: Our work with idealized images has successfully demonstrated our basic approach to analyzing shape. With our theoretical framework and implemented tools we will now attempt to demonstrate the efficacy of our approach on real images. Initially we will focus on detecting human bodies and faces within real images.

References:

[1] Harry Blum. Biological Shape and Visual Science. *Journal of Theoretical Biology*, 38:205-287, 1973.

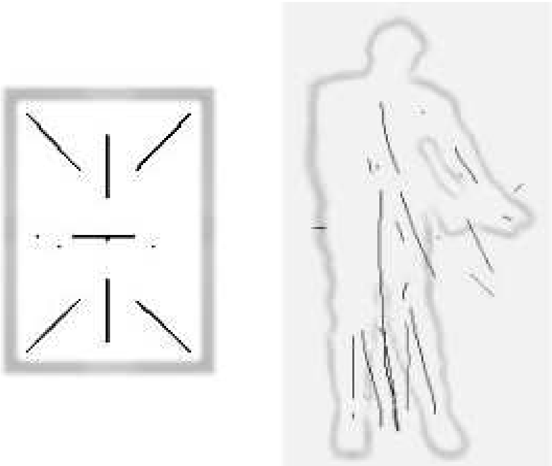

Figure 1: Left: Salient Medial Axes of a Simple Rectangle. Right: Simple Axes Corresponding to Near Parallel Edges of an Idealized Human Silhouette.

[2] David Marr. Vision. W.H. Freeman and Co., New York, 1982.

[3] Bryan S. Morse, Stephen M. Pizer, Derek T. Puff and Chenwei Gu. Zoom-Invariant Vision of Figural Shape: Effects on Cores of Image Disturbances. *Computer Vision and Image Understanding*, 69(1):72-86, 1998.

[5] Marek Teichmann and Seth Teller. Assisted Articulation of Closed Polygonal Models. *MIT Computer Graphics Group, graphics.lcs.mit.edu*.

[6] Jules Bloomenthal. Skeletal Design of Natural Forms. PhD thesis, CS Dept., University of Calgary, Alberta, Canada, 1995.

[7] Michael F. Kelly and Martin D. Levine. Edge Detection Within an Annular Sampling Framework. *TR-CIM-95-4*, Centre for Intelligent Machines, McGill Univ., 1995.

[8] Song Chun Zhu and A. L. Yuille. FORMS: A Flexible Object Recognition and Modeling System. *International Journal of Computer Vision*, 20(3):187-212, 1996.

[9] Purves, Orians, Heller and Sadava. Life: The Science of Biology, 5th edition. W.H. Freeman, Utah, 1998.

[10] Mourad Zerroug and Ramakant Nevatia. Segmentation and 3-D Recovery of Curved-Axis Generalized Cylinders from an Intensity Image. *International Conference on Pattern Recognition*, A:678-681, 1994.