MVA'90 IAPR Workshop on Machine Vision Applications Nov. 28-30,1990, Tokyo

Recognition of Human Motion based on Interpretation of 2D Pattern Deformation

Keiko Shiraishi, Mutsuhiro Terauchi and Kenji Onaga

Faculty of Engineering Hiroshima University

Abstract

This paper presents an approach to recognize human motion from single viewed image sequence in order to realize a new man-computer interface. The human body is assumed to be linked model of rigid segments which are approximated by ellipsoids with three different length of diameter. The reconstruction of human motion is mainly performed by interpretation of 2D pattern deformation as 3D motion. We show the effectiveness of utilizing changes of curvature and discontinuity on the contour of 2D pattern for human motion reconstruction.

1. Introduction

In the previous computer-work environments between human and computer the transmission of information has been based on text representation on the computer terminals. Recently it is evolved to be more interactive using visual and sound effects. Yet the transmission is almost uni-directional as ever, from computer to user. So the new input device for computer is desired to be developed as a natural interface between human and computer in the next generation.

One of the most useful information is auditory one, which has been well studied and begins to be applied for real tasks especially in text input. The exchange of information using auditory media uses almost linguistic data which has the meaning of words or sentences themselves. It follows that the linguistic data exchange can be replaced by text input/output using computer keyboard and display.

By the way the visual information is the most dominant in human senses. Under the condition that it is too noisy or that they can not understand opponents' language, gestures or movements of their hands are effective methods for understanding one another. They are called as body language, and often used in real communication environment. Thus it is strongly desired for human computer interface to equip computers with ability of visual pattern recognition.

In order to reconstruct and understand the human posture and motion from sequential images, it is necessary (1) to define the shape of human model, (2) to extract valid features from each image, (3) to reconstruct 3D posture of human body from them and (4) to give consistent interpretation of the motion using knowledge about human motion and link model. There are however a lot of problems in each step, for example complicatedness of shape of human body in step (1), occlusion of segments in (2),(3) or hierarchical representation of motion with valid meaning in (4).

In the previous studies, Badler and O'Rourke defined motion representation using simulated image of human motion generated by computer [1], and analized sequential images by using constraint propagation operator [2]. Akita represented body segment as general cone, and improved a matching method between features in two different images in which ordered set of feature values extracted from a small window area, called window code, was computed [3]. Etoh et.al. expanded the method, and extracted link axis using extremal curvature points as vertices of polygon (called ribbon) in order to make it of confidence [4]. They are summarized that there are two processes, one is fitting an axis into separated single link pattern on an image, and another is interpretation of 2D linked axes as 3D posture of human body.

In this paper we assume that segmented parts of human body can be approximated by ellipsoid and is rigid, and analyze the image sequence on the viewpoint that 3D motion of the rigid ellipsoid causes 2D pattern deformation in which the curvature of pattern contour changes according to 3D motion in order to reconstruct 3D posture and motion of human body.

2. Analysis of image sequence

The purpose of our research is first to reconstruct 3D posture and motion of human body from projected image sequence, secondly to give an interpretation of overall reconstructed motion and last to predict the future motion from given image sequence.

The image sequence of human motion is given as the set of still images. After a image sequence is given, the projected contours are extracted from each image, then they are segmented into the parts of human body which amount to rigid links. We consider the each parts as individual object and reconstruct its 3D posture. Next we connect the body parts based on the connectivity and the metrics of given human body model in order to generate the human like posture. The human posture is reconstructed from each image, then motion can be derived from transition between two neighboring posture image. In addition to the method we can infer the body motion directly from image sequence by interpreting the deformation of 2D projected contour as 3D movement.

If there is an assumption that the object is rigid and continuously smooth so that the projected contour is differentiable, then the deformation of 2D pattern can be interpreted as 3D motion. Furthermore when there exists discontinuous point it is considered that the object is occluded by another object at the point.

3. Representation of human body

When we reconstruct the 3D motion and posture from image sequence, it is necessary to define the model which suits for human object. In the present paper we introduce the human model which consists of rigid ellipsoids (Fig.1) with different metrics, where each link amounts to real part, e.g. head, arm, trunk, leg etc. Three diameters of ellipsoid are given for all links in prior. The human body model has also the knowledge about joint position, connectivity and geometical relationship between links.

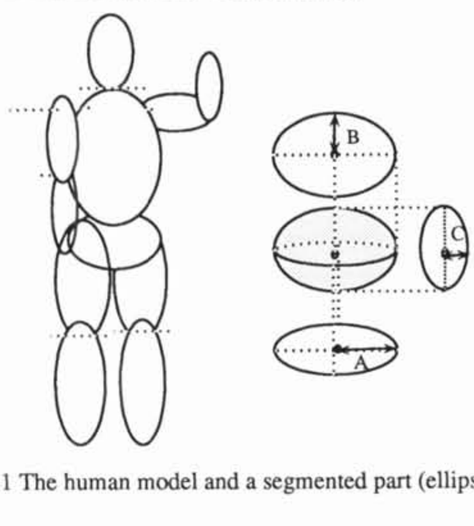

Fig.1 The human model and a segmented part (ellipsoid).

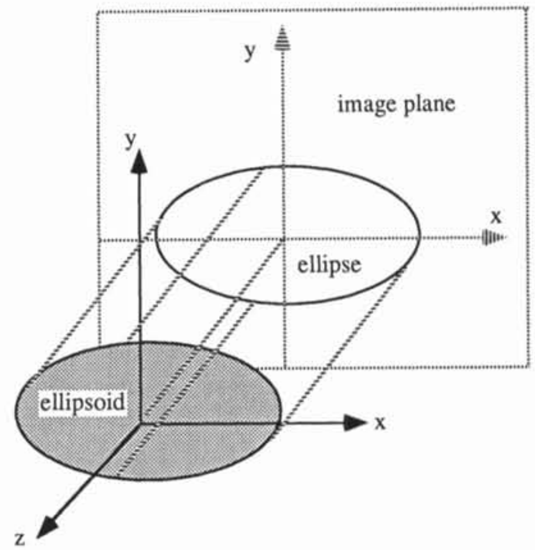

Fig.2 The orthographic projection of an ellipsoid.

The advantages of introducing the ellipsoid model are that it only has positive curvature on the surface and that the 3D motion of it is observable from its projection except for the translation along viewing axis. However there is a problem that the interpretation of the posture of ellipsoid is not obtained uniquely.

4. Geometry of projection

In general a 3D scene is projected onto the an image by perspective projection through the lens system as human eyes or artificial camera system.But when the distance between the viewpoint and the object point is enough long comparing to the object size, the projection can be approximated as the orthographic one. In this paper we introduce orthographic projection to formulate the relation between 3D object and projected pattern by following reasons,

* There is no deformation according to the distance between viewpoint and object point, i.e. the geometry can be simplified.

* The motion can be formulated simply.

We assume that the link element of the human model are approximated as ellipsoid. Then it is necessary to formulate the relationship between the posture of ellipsoid in 3D space and the projected contour on the 2D image (Fig.2).

Furthermore our attention concentrates not on the position of ellipsoid but on posture so that we can assume the origin of the coordinate axes to be on the centroid of the ellipsoid. In this case the equation of the ellipsoid can be written as follows neglecting first order of equation :

$$a_1x^2+b_2y^2+c_1z^2+d_1xy+e_1yz+f_1xz+g_1=0 \qquad (1).$$

The expected shape of the projected contour is an elipse, so we aim to derive the quadratic equations from the equations as follows:

$$Ax^4+(By)x^3+(Cy^2+D)x^2+(Ey^3+Fy)x+(Gy^4+Hy^2+I)=0 \quad (2).$$

where the coefficients A,B,C,D,E,F,G,H and I are represented as follows:

$$A=c_1^2(e_1^2-4a_1c_1)^2(a_1f_1^2-b_1e_1^2+d_1e_1f_1)^2\ ,$$

$$B=-c_1(e_1^2-4a_1c_1)(2b_1e_1-d_1f_1)(2a_1f_1-d_1e_1)(a_1f_1^2-b_1e_1^2+d_1e_1f_1),$$

$$\begin{aligned}C=&-2c_1(-(e_1^3(-(d_1+c_1e_1f_1))+b_1^2(c_1e_1(2d_1^2+e_1^2f_1^2))\\&+b_1^3(2e_1^3-4c_1^2e_1^3))+b_1(-2d_1^2f_1(c_1d_1-2c_1^2e_1f_1\\&+e_1f_1))))+a_1^3(2f_1^4(-1+2c_1^2)(-4b_1c_1+f_1^2))+\\&a_1^2(-(f_1^2(c_1f_1^2(2d_1^2+e_1^2f_1^2)+b_1^2(-16c_1e_1^2(-1+\\&2c_1^2))+b_1(2e_1f_1(-4c_1d_1-e_1f_1+2c_1^2e_1f_1)))))+\\&a_1(2e_1(-(d_1^2f_1^3(-(c_1d_1)-e_1f_1+2c_1^2e_1f_1))+b_1^3\\&(-4c_1e_1^3(-1+2c_1^2))+b_1^2(-(e_1^2f_1(-4c_1d_1-e_1\\&+2e_1f_1)))+b_1(e_1f_1^2(-6c_1d_1^2+8c_1^3d_1^2-2d_1e_1f_1\\&+c_1e_1^2f_1^2))))\ ,\end{aligned}$$

$$D=-c_1g_1(e_1^2-4a_1c_1)\{8c_1(b_1e_1^2-a_1f_1^2)(a_1f_1^2-b_1e_1^2+d_1e_1f_1)-e_1^2(2b_1e_1-d_1f_1)^2\}\ ,$$

$$E=c_1(f_1^2-4b_1c_1)(2b_1e_1-d_1f_1)(2a_1f_1-d_1e_1)(b_1e_1^2-a_1f_1^2+d_1e_1f_1),$$

$$F=c_1g_1(2b_1e_1-d_1f_1)(2a_1f_1-d_1e_1)\{2c_1(b_1e_1^2-a_1f_1^2)-f_1^2(e_1^2-4a_1c_1)\}\ ,$$

$$G=c_1^2(f_1^2-4b_1c_1)^2(b_1e_1^2-a_1f_1^2+d_1e_1f_1)^2\ ,$$

$$H=-c_1g_1(f_1^2-4b_1c_1)\{8c_1(b_1e_1^2-a_1f_1^2)(b_1e_1^2-a_1f_1^2+d_1e_1f_1)-f_1^2(2a_1f_1-d_1e_1^2)^2\},$$

$$I=16c_1^4g_1^2(b_1e_1^2-a_1f_1^2)^2.$$

To derive ellipse contour equation, we factorize the biquadratic equation (2) into the pair of quadratic equations. As the result, three candidates are obtained as following,

$$\{x^2-lx+(l+m+n)(l-m-n)/4\}\{x^2+lx+(l+m-n)(l-m-n)/4\}=0\ ,$$
$$\{x^2-mx+(l+m+n)(l+m-n)/4\}\{x^2+mx+(l-m-n)(l-m+n)/4\}=0\ ,$$
$$\{x^2-nx+(l+m+n)(l-m+n)/4\}\{x^2+mx+(l-m-n)(l+m-n)/4\}=0,$$

where l, m, n are the resolutions of cubic equation.like as follows,

$$l^2=\frac{s}{3}+\frac{\frac{-s^2}{9}+\frac{t}{3}}{\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}+\left(\left(\frac{-s^2}{9}+\frac{t}{3}\right)^3+\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}\right)^2\right)^{\frac{1}{2}}\right)^{\frac{1}{3}}}-\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}+\left(\left(\frac{-s^2}{9}+\frac{t}{3}\right)^3+\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}\right)^2\right)^{\frac{1}{2}}\right)^{\frac{1}{3}},$$

$$m^2=\frac{s}{3}-\frac{\sqrt{3}\,\dfrac{\frac{-s^2}{9}+\frac{t}{3}}{\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}+\sqrt{\left(\frac{-s^2}{9}+\frac{t}{3}\right)^3+\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}\right)^2}\right)^{\frac{1}{3}}}}{2}-\frac{\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}+\sqrt{\left(\frac{-s^2}{9}+\frac{t}{3}\right)^3+\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}\right)^2}\right)^{\frac{1}{3}}}{2}-\frac{\frac{-s^2}{9}+\frac{t}{3}}{\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}+\sqrt{\left(\frac{-s^2}{9}+\frac{t}{3}\right)^3+\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}\right)^2}\right)^{\frac{1}{3}}}+\frac{\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}+\sqrt{\left(\frac{-s^2}{9}+\frac{t}{3}\right)^3+\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}\right)^2}\right)^{\frac{1}{3}}}{2},$$

$$n^2=\frac{s}{3}+\frac{\sqrt{3}\,\dfrac{\frac{-s^2}{9}+\frac{t}{3}}{\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}+\sqrt{\left(\frac{-s^2}{9}+\frac{t}{3}\right)^3+\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}\right)^2}\right)^{\frac{1}{3}}}}{2}+\frac{\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}+\sqrt{\left(\frac{-s^2}{9}+\frac{t}{3}\right)^3+\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}\right)^2}\right)^{\frac{1}{3}}}{2}-\frac{\frac{-s^2}{9}+\frac{t}{3}}{\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}+\sqrt{\left(\frac{-s^2}{9}+\frac{t}{3}\right)^3+\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}\right)^2}\right)^{\frac{1}{3}}}+\frac{\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}+\sqrt{\left(\frac{-s^2}{9}+\frac{t}{3}\right)^3+\left(\frac{-s^3}{27}+\frac{st}{6}-\frac{u}{2}\right)^2}\right)^{\frac{1}{3}}}{2},$$

where the coefficients s,t,and u are expressed as follows ;

$$s=\frac{3B^2y^2}{4A^2}-\frac{2(D+Cy^2)}{A},$$

$$t=\left(\frac{-3B^2y^2}{8A^2}+\frac{D+Cy^2}{A}\right)^2-4\left(\frac{-3B^4y^4}{256A^3}+\frac{B^2y^2(D+Cy^2)}{16A^2}-\frac{By(Fy+Ey^3)}{4A^2}+\frac{I\ Hy^2+Gy^4}{A}\right),$$

$$u=\frac{B^3y^3}{8A^3}-\frac{By(D+Cy^2)}{2A^2}+\left(\frac{Fy+Ey^3}{A}\right)^2.$$

$$\omega^3+s\omega^2+t\omega+u=0$$

Then we can calculate ellipse contour in 2D image and the coefficients of projected contour from these relationships. It follows that it is possible to reconstruct the posture of the ellipsoid in the 3D space from projected contour.

Table.1 Relationship between motion of an ellipsoid and features of the projected contour.

motion in 3D space		curvature extreme		link axis	
		location	curvature	terminal point	length
translation	along Z axis	no	no	no	no
	other direction	translation	no	translation	no
rotation	around Z axis	rotation	no	rotation	no
	other rotation	yes (uncertain)	yes	movement on elliptic trajectory	prolonged or shortened

5. The property of feature in 2D image

In order to segment overall contour into ellipse shaped parts, it is useful to utilize the features extracted from the contour. The representative of them are as follows,

1) discontinuity point
2) curvature extreme

These are useful features for dividing the overall contour into links.

Leyton proposed that closed smooth planar curve is roughly represented by curvature extreme [5]. Fundamentally the projected contour of human body is considered to be closed smooth planar curve. So the contour can be roughly represented by curvature extreme points. In our human body model each link is approximated by ellipsoid, which has only the positive curvature extreme, so that negative curvature extreme or discontinuity point appears around the joint point or on the intersection of occluding boundary.

5.1 discontinuity point

When human body is projected onto 2D plane,there generally exists occluding area. This is caused by overlapping of two or more object parts. These contour intersection points on the image become discontinuous. point. The overall overlapped contour should be divided into each part (ellipse) to recognize the deformation (shape) of each ellipse. Then we introduce the method to generate subjective contour which is usually applied in psychophysical simulation of human visual phenomena. In the method the terminal point, which is given as discontinuous point here, is grown with good-continuation hypothesis.

5.2 Curvature extreme

If a disk is rotated around an arbitrary diameter axis in the 3D space, their extreme curvature of projected contour in the image changes in response to the rotation (Table.1) . From this point of view, it is considered that the object rotation in 3D space can be computed from the change of the extreme curvature of projected contour.

6. Conclusion

In this paper we consider the framework in which we interpret deformation of projected pattern in image sequence as 3D motion.

We formulate here the relationship between 3D posture of a ellipsoid and its projected ellipse contour. It is necessary that we elicidate the limitation of this method and combine with quantity method. Furthermore we want to predict the next motion in the image sequence and to interpret what motion human do.

Acknowledgment

The work presented in this paper was supported in part by Kumahira Security Foundation. We would like to express our appreciation to Dr. Ito and Dr. Tsuji for their valuable suggestions.

References

[1] N.I. Badler and S.W. Smoliar, " Digital representation of human movement, " Ass. Comput. Surv. 11, pp19-38 (1979).

[2] J. O'Rourke and N. I. Badler, " Mode-based image analysis of human motion using constraint propagation, " IEEE Trans. Pattern Anal. Mach Intell., PAMI-2, pp522-536 (1980).

[3] K. Akita, " Image sequense analysis of real world human motion, " Pattern Recognition, Vol.17, No.1, pp73-83 (1981).

[4] M. Etoh , A. Tomono and F. Kishino, " Stereo Based Description by Generalized Cylinder Complexes from Occluding Contours, " Trans. IEICE, Vol.J73-D-2, pp1402-1412 (1990).

[5] M. Leyton, " A process-grammar for shape, " Artificial Intelligennce, Vol.34, pp213-247 (1988).