Recovery of 3D Models from Sequences of Images

Nan Weng

Supervisor: Dr. Herbert Yang

Abstract

The recovery of 3D model from 2D image is useful in medical image analysis, visual communication, education, entertainment and etc. Image sequences got from the relative motion between the object and the camera or scanner contain more 3D information than a single image. There are several methods to reconstruct the 3D model from image sequences. These methods can be divided into two categories, the feature correspondence approach and the optical flow approach. The feature correspondence approach has some limitations. In this project, the optical flow approach was studied. Optical flow is the apparent motion in the image plane produced by the projection of the 3D real motion into 2D images. The 3D real motion of an object can be recovered from the optical flow field by adding constraints. The 3D velocity vectors contain the information of the surface of the object. By extracting the information of the 3D motion, it is possible to get more accurate 3D model of the object. In this project, a new method using optical flow information obtained from the image sequences is proposed to reconstruct 3D models. Experiments are done using synthetic and real image sequences. From the experiments, one can see that this method is feasible, but improvement is still needed.

1 Introduction

1.1 Motivation

In the last 10 years, the problem of reconstructing and representing three-dimensional shapes has attracted a remarkable interest among the researchers in the fields of computer vision and computer graphics. A trend of recent research in computer vision is to develop new algorithms to recover 3D model from 2D images. There are increasing demands for the developments in model-based coding, stereoscopic display, and graphics-generated virtual environments that require three dimensional modeling of various objects. Application areas for these include visual communications, education and entertainment.

Another important area which requires the 3D model recovery very much is in medical applications. Doctors can get different kinds of images, like X-ray Computed Tomography (X-CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) and Ultrasound Images. Most of the 2D medical images are cross sections of objects. The reason is that it is often necessary to acquire data from methods that essentially “tear” data apart (or acquire the data one piece at a time) in order to be able to view what is inside the object with complicated structure, like the brain and most organs inside the body. Also it is usually quite difficult to get a photograph of the object as a whole, like the fetus inside the mother’s uterus. The sequence of cross sections of an object contains the shape information of the object, from the outer most surface to the inner most surface. If the 3D shape of an object could be visualized from the slices, it would be very helpful for doctors to analyze the object. So, the recovery of the 3D shape from cross sections, the process which is different from the recovery of 3D models from usual photographs, is a very important research topic.

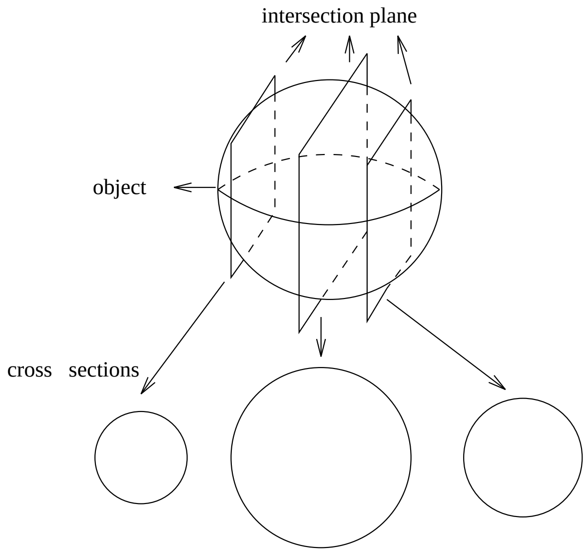


Figure 1: Cross sections of an object.

Fig.1 shows the mechanism of acquiring an image sequence of cross sections. In the main project of CMPT 880, a new method which uses a sequence of ultrasound images to reconstruct the 3D model of a fetus has been implemented.

1.2 The recovery of 3D model from an image sequence

Among the available shape cues in a 2D image, texture gradient, shading, surface contour and perspective distortion have been studied. A single image only provides a sparse dataset. Mathematically speaking, the information available in a single 2D image is always ambiguous. Only under certain regularity assumptions can local evidence about the 3D shape of an object be derived [3]. So using an image sequence generated due to translation or rotation or both of an object will provide more depth information for 3D model recovery process.

The methods of recovering 3D models from 2D image sequences can be divided into two categories, the feature correspondence approach and the optical flow approach.

There are many algorithms belong to the feature correspondence approach. For example, the one proposed in [2] is based on boundary detection and derivative of contour movement. Due to the mechanism of acquiring the photographs, the method suffers from surface discontinuities caused by occlusions. Sun and *et al* [11] propose a method to reconstruct the kidney using the ultrasound images of kidney, based on an elliptic deformable model and interpolations with sparse data set, which is not suitable for the recovery of organs with more complicated structures and shapes. In [12], a hierarchical spline surface is proposed as the representation scheme, which is only applied to low order surfaces.

In general, the feature correspondence approach depends on geometric primitives, such as point, line, or volumetric feature. So, there are many difficulties and limitations encountered with this approach, e.g. complicity of the surface, noise and the choice of feature representations.

1.3 The concept of optical flow

Optical flow is the apparent motion in the image plane produced by the projection of the real motion of object in 3D world onto 2D images. In this project, the focus is on the optical flow approach.

Actually, optical flow is a measure of the change of image brightness in an image sequence and is commonly used as an approximation of the velocity field, which is the perspective projection of the three-dimensional real velocity on the image plane. Optical flow results from the relative motion between the camera and the objects in the scene.

There is no direct way to calculate the real velocity from an image sequence. The reason to use optical flow as an approximation of the velocity field is that, in two specific situations, these two fields coincide. One is under isotropic illumination, provided that the optical system is calibrated, the object has Lambertian surfaces, patterns are locally rigid and the motion is contained in a plane parallel to the image plane (the rotational components are only around the Z-axis). Another is under uniform light source, the object is composed of planar patches and has Lambertian surfaces, patterns are locally rigid and the motion is translational and is on plane parallel to the image plane. Even if these conditions are not strictly satisfied, in many applications, optical flow is a good approximation of the velocity field.

We can get much useful information from the optical flow obtained from an image sequence, like the approximate motion direction, the relative velocity between objects moving with different velocities. So, the calculation of optical flow can be perceived as an early processing step for some spatiotemporal approaches. If additional constrains are incorporated, the real 2D velocity and, hence, the 3D velocity of the real motion can be recovered from the optical flow vectors.

1.4 The calculation of optical flow

The gradient-based approach estimates the optical flow using the derivatives of image brightness. called gradient-based methods. The gradient-based constraint is derived based on the assumption that in the immediate neighborhood of a pixel, the optical flow field is smooth, i.e, the Taylor series expansion can be applied.

Based on this assumption, one can get the optical flow constraint equation:

$$\bigtriangledown \vec{f}(s) \bullet \vec{\omega_s} + f_t = 0$$

, where s=(x,y,t) is a point in the image at time t. (x,y) designates the 2D spatial image coordinates and t the time axis. f(s) is the observed intensity function. $\bigtriangledown \vec{f}(s)$ is the spatial image gradient $(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$ and f_t stands for $\frac{\partial f}{\partial t}$. $\vec{\omega_s} = (u_s, v_s)^T$, where $u_s = \frac{dx}{dt}(s), v_s = \frac{dy}{dt}(s)$. We can see that $\vec{\omega_s}$ is the optical flow.

This equation is also called the standard motion constraint equation proposed by Horn and Schunck. The constraint equation itself does not provide a means for calculating optical flow. The equation only constrains the values of u_s and v_s to lie on a line in the gradient space.

As the authors of [1] said, the estimation of dense velocity fields is basically an ill-posed problem, mainly because the data only partially constrain the solution. So, we have to add additional constraints before we can calculate the optical flow.

There are several ways to incorporate additional constraints, least-square being one of them. The least-square method has several variation, e.g. using two consecutive images or using several pairs of consecutive images.

In this project, the method of least-square using two consecutive images was implemented. A relaxation algorithm was applied. Iteration continues until an error measure is minimum.

Fig.1 shows two circles, the front one is undergoing dilation and the occluded one is moving to the lower right corner. Fig.2 is the optical flow field obtained from two consecutive images of the moving circles. The direction of the motion can be seen very clearly. Because of the limitation of the method, the error is propagated so that the points which are not moving can also have non-zero optical flow values. In the following implementation, this kind of error is avoided.

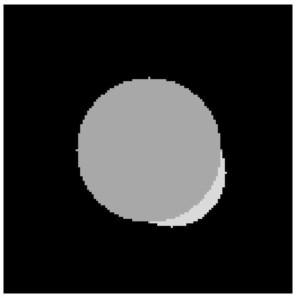


Figure 2: Two circles undergoing dilation and translation.

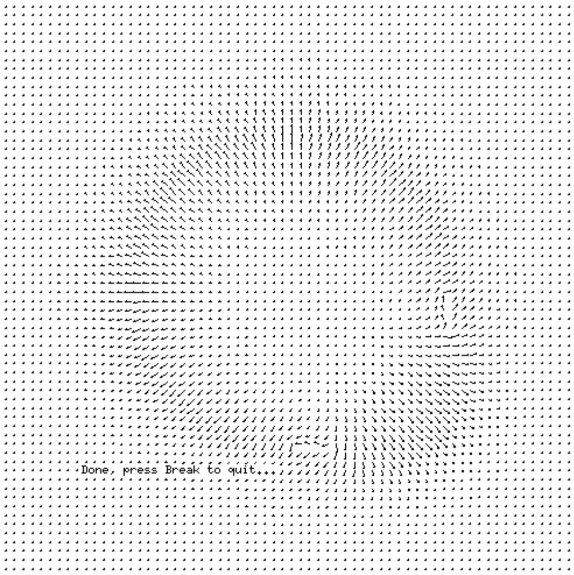


Figure 3: The optical flow field of the two circles in Fig.2.

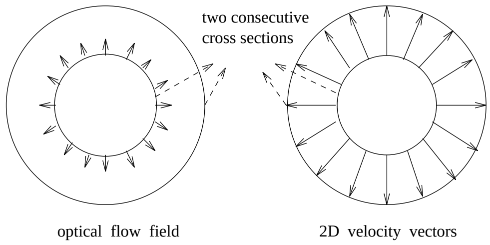


Figure 4: The relation between the optical flow and the velocity.

2 Theory and implementation

Two dimensional motion carries information about the third dimensional, which can be used to get the depth information and the surface orientation which are necessary for the recovery of a 3D model [8]. Discontinuities in optical flow field have been considered in previous work and no geometric representations of features are required. Hence, the optical flow methods may overcome problems in other algorithms which depend on the feature extraction process.

So far, there are not many methods proposed using the information of optical flow directly, since there are still some limitations with the optical flow approach as can be seen later in this paper. The recovery of 3D motion from 2D optical flow field and then using the 3D motion information to recover the surface of the object is not trivial. Shu and Shi [13] propose an algorithm to recover the Nth order surface structure using the unified optical flow field. But for most objects, especially in medical application, an Nth order surface is not enough because of discontinuities.

2.1 Optical flow and 3D motion

The recovery of 3D motion from 2D information is a very challenging research problem. Different situations need different constraints and assumptions. In order to solve more complicated problem, more constraints need to be added.

Most modern medical imaging systems, like Ultrasound Image, MRI, PET, X-CT, etc. , provide the cross sectional views of an object. We can get a sequence of cross sections of an object by moving the scanner along an axis of the object. This is the relative motion between the object and the scanner, and is equivalent to images generated with a stationary scanner and a moving object. Because of the equivalence, the use of optical flow is employed. By calculating the apparent motion of the pixels between each pair of the consecutive images, we can get the 2D optical flow vectors.

The 2D optical flow vectors stand for the direction of 2D velocities projected from the 3D motion. From the cross sections, the contour in each image can be computed using a simple edge detector. Then the 2D velocity vectors in terms of pixels can be calculated easily. From Fig.4, we can see the relation between the optical flow and the velocity values.

Because of the precision of the scanner, the displacement between two slices (cross sections) is fixed, then the velocity component along the moving axis in terms of pixels is known. Suppose all the projections are orthogonal, then the 3D velocity vectors can be calculated, and hence, the 3D motion is recovered.

2.2 Calculation of the surface normal

In Fig.3, one can see that the 2D optical flow vectors in each slice are vertical to the boundary of the object. It is always true for the optical flow got from the two consecutive cross sections of the object. Then the tangent vector of the contour can be calculated from the 2D optical flow.

Intuitively, the direction of 3D motion coincides with the tangent of the surface point of the object, i.e. the 3D velocity vectors always vertical to the normal of the surface. Because the surface normal

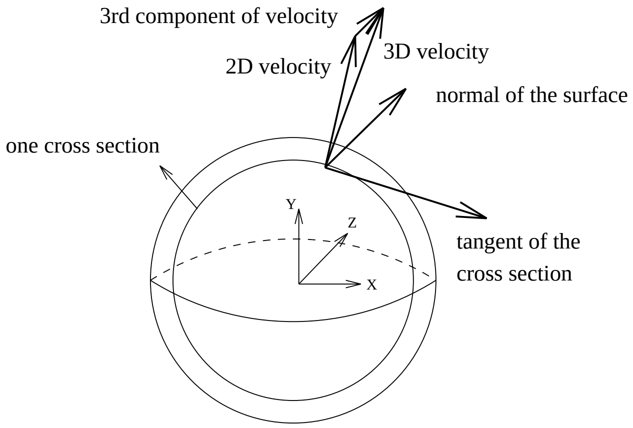


Figure 5: The method to calculate the normal of the surface.

is also vertical to the tangent vector of the cross section. So, the normal of the surface point is the cross product of the tangent vector of the contour of the cross section and the 3D velocity vector of that point.

Fig.5 illustrates the way to calculate the normal of the surface. The key point of the method is the recovery of the 3D velocity vector from the 2D optical flow field. The medical imaging system makes the problem much simpler because the third component of the velocity vectors are fixed and can be estimated in terms of pixels accurately from the object size.

2.3 Algorithm

The proposed algorithm can be summarized as follows:

- get an image sequence,
- calculate the 2D optical flow of a pair of images,
- calculate the 2D velocity vectors from the optical flow field,
- add the 3rd component of the velocity,
- calculate the normal of the surface at each point,
- use the technique of graphics, ray-tracing, to visualize the recovered surface .

3 Experiments and analysis:

3.1 Data set

In the implementation, the volume data set were used. Each point in the volume data set is called a voxel instead of pixel in 2D data. The coordinate is (X, Y, Z), X and Y stand for the position of the voxel in the slice, and Z is the sequence number of the slice on the time axis along the scanning direction, i.e the third direction of the relative motion. An intensity is associated with each voxel.

If the data is obtained from rotational scanner, the same data set as that got from a translational scanner can be computed by transformation and interpolation.

In the experiments, both synthesized and real images were employed. The synthesized image of a sphere (128 X 128 X 98), a diamond (128 X 128 X 121) and an ellipsoid (64 X 64 X 61) were used. As to the real image sequence, an ultrasound image sequence of a fetus obtained by scanning the fetus

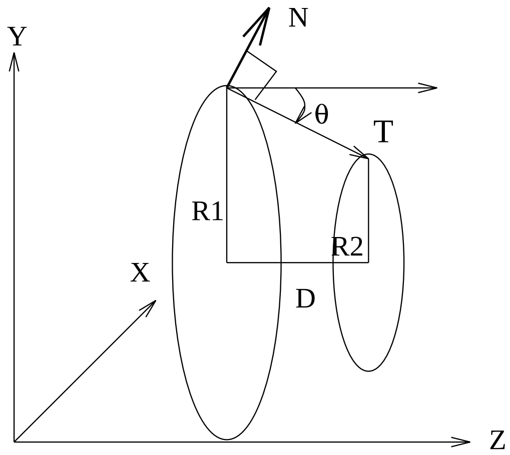


Figure 6: The theoretical way of calculating the surface normal.

from the head to the bottom was used. Transformations were done to get another two data sets, one of which scans from the frontal side to the back and the other from the left side to the right side of the fetus.

Different from the other image acquiring mechanisms, the medical images of an object are very dense, and can be made even denser by interpolation.

3.2 Verification

The theory of calculating the normal of the surface point is verified through the experiments with the synthetic images of a sphere. Fig.6 shows the method of calculating the normal of the surface from two consecutive slices of the sphere. $\vec{N}$ is the unit normal at the point of interest on the surface. R1 and R2 are the radii of the two circles which are the two consecutive cross sections of the sphere. D is the distance between the two slices and is known. θ is the angle between the z axis and the unit tangent vector. $\vec{T}$ is orthogonal to $\vec{N}$. Because R1, R2 and D are known, then $\tan\theta = \frac{R1-R2}{D}$ can be calculated. The cross product of $\vec{N}$ and z axis is equal to $sin\theta$, i.e, $\vec{N} \cdot \vec{Z} = \sin\theta$. Hence, the accurate values of $\sin\theta$ can be calculated theoretically. After the normal at each surface point was calculated using the optical flow approach as described in section 2, the cross products of the normal and the z axis were compared to the theoretical values. The experimental results coincide with the theoretical values within 2%. This verifies the feasibility of the proposed method.

Another test was performed by varying the third component of the velocity, i. e. D. If the value is less than what it should be, intuitively, the recovered sphere will look flatter than the original one. Since the radius of the sphere is 50 pixels, and the number of slices is nearly 97, so the third component of the velocity $\vec{V_z}$ is 1 in terms of pixel. Comparing Fig.7.1 with Fig.7.2, one could verify the correctness of the implementation qualitatively, the one with a smaller value of $\vec{V_z}$ looks brighter and flatter.

3.3 Experimental results

After computing the normal at each surface point, the shading value was assigned by assuming the surface is Lambertian, i.e. the intensity is proportional to the cosine of the normal and the direction of the light source.

Because of the page limitations, only selected experimental results are included in this paper. In Fig7.1, the missing points in the center of the sphere are due to the sparseness of the data set.

In the implementation, the gradient-based method to compute the optical flow is used. There are errors with the method, and the error propagates to the points which are not moving. So, interleaving of the different slices, an edge detector was used to extract the information along the edges of each slice. Because of the displacement between the two slices, there are discontinuities on the recovered surface. Interpolation is needed to complete the surface. Increasing the density of the data set may be needed.

The proposed method was also applied to an ultrasound image sequence of a fetus. Fig.8.1 is the result of running the data set scanning from the left to the right sides of the fetus. Fig.8.2 is the result of volume rendering of the same fetus data for comparison. Although the result is not so satisfying, one can tell the shape of the fetus, a big head with two hands in front of his body.

Noise is a serious problem in ultrasound images. Because an edge detector was used during the implementation, the threshold of the edge detector would effect the results. Fig.9.1 and Fig.9.2 are results computed from the data set which scans from the frontal side to back with different thresholds. When the threshold was too small, the noise made the reconstructed image quite messy. While it was too high, a lot of real data was lost. When the threshold was equal to 40, the results appears reasonable.

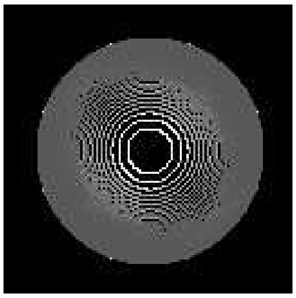
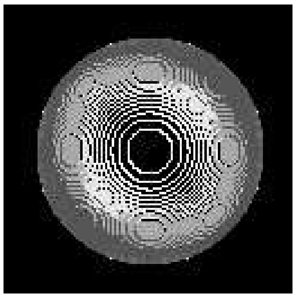
1  2 

Figure 7: 1:$\vec{V}_z = 1$, 2: $\vec{V}_z = 0.2$.

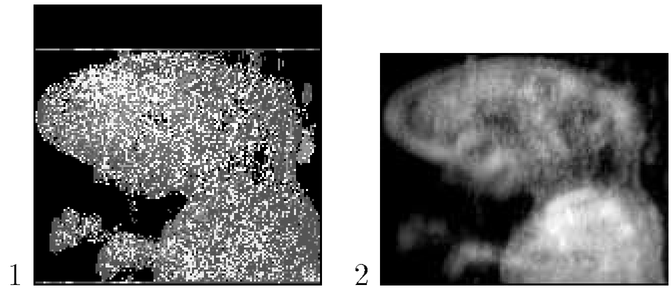


Figure 8: 1: 3d surface of the fetus, 2: volume rendering result.

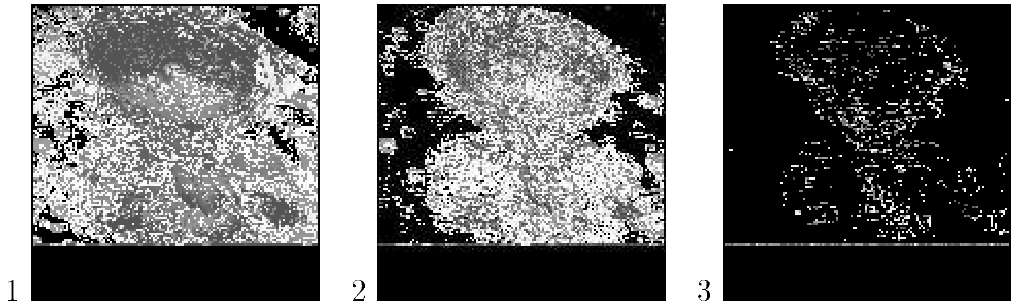


Figure 9: 1: threshold=20, 2: threshold=40, 3: threshold=80.

3.4 The problems with the proposed method

There are some difficulties and limitations with the optical flow approach. First of all, the optical flow requires a substantial amount of calculations. So, it is not suitable for real time processing.

The value of the 2D optical flow will affect the recovery of shape of the object. The algorithm for reconstructing the 3D model assumes that the image flow velocity field has already been computed and is readily available as an intrinsic image. So far, there is no perfect method to calculate the optical flow since the problem is only partially constrained. The added constraints are usually not suitable to any situation. So it is difficult to obtain an accurate value of the optical flow. Furthermore, noise as well as large displacements will also affect the results.

4 Conclusion and future work

In the project, a new method was implemented to recover the 3D model from an image sequences using optical flow information, combining the techniques in computer vision and computer graphics.

The experiments showed that this method is feasible, and accurate in calculating the normal of the surface. But the implementation still needs to be improved to address the problem of noise and incomplete interpolation.

The results of real images are not satisfying. The main reason may be the sensitivity of the edge detect to noise. Points which are not on the real contours may distort the reconstructed surface. Currently, a new method of abstracting more accurate and succinct contours using connected components and mathematical morphology is being investigated. Also, synthesized data set with high resolution will be generated to address the problem of the missing points.

References

[1] HEITZ, F. AND BOUTHEMY, P. Multimodal Estimation of discontinuous optical flow using Markov random fields. *IEEE Trans. Pattern Anal. Machine Intell. 15*, 12 (Dec. 1993), 1217-1232.

[2] ZHENG, J. Y. Acquiring 3-D Models from sequences of contours. *IEEE Trans. Pattern Anal. Machine Intell. 16*, 2 (Feb. 1994), 163-178.

[3] POLI, R., COPPINI, G. AND VALLI, G. Recovery of 3D closed surfaces from sparse data. *CVGIP: Image Understanding 60*, 1 (July 1994), 1-25.

[4] LEVOY, M. Volume Rendering: Display of surfaces from volume data. *IEEE Computer Graphics and Applications 8*, 3 (May 1988), 29-37.

[5] PRAZDNY, K. On the information in optical flows. *Computer Vision, Graphics, and Image Processing 22*, 2 (May 1983), 239-259.

[6] TERZOPOULOS, D. Multilevel computational processes for visual surface reconstruction. *Computer Vision, Graphics, and Image Processing 24*, 1983, 52-96.

[7] PEI, S. AND LIOU, L. Finding the motion, position and orientation of a planar patch in 3D space from scaled-orthographic projection. *Pattern Recognition 27*, 1 (Jan. 1994), 9-25.

[8] MARTIN, W. N. AND AGGARWAL, J. K. *Motion Understanding: Robot and Human Vision.* Kluwer Academic Publishers, 1988.

[9] BEN-TZVI, D., DEL BIMBO, A. AND NESI, P. Optical flow from constraint lines parameterization. *Pattern Recognition 26*, 10 (Oct. 1993), 1549-1561.

[10] XU, J. *Image Processing and Analysis.* Science Press, China, 1992.

[11] SUN, Y-N., LEE, J-S., CHANG, J-C. AND YAO, W-J. Three-dimensional reconstruction of kidney from ultrasound images. In *Proc. of the IEEE Workshop on Biomedical Image Analysis*, June 1994, Seattle, Washington, pp. 43-49.

[12] WANG, Y. F. On 3D model construction by fusing heterogeneous sensor data. *CVGIP: Image Understanding 60*, 2 (Sept. 1994), 210-229.

[13] SHU, C. Q. AND SHI, Y. Q. Direct recovering of Nth order surface structure using unified optical flow field. *Pattern Recognition 26*, 8 (Aug. 1993), 1137-1148.