Surface Shape Reconstruction of an Undulating Transparent Object

Hiroshi MURASE

NTT Basic Research Laboratories,
3-9-11, Midori-cho, Musashino-shi, Tokyo 180, JAPAN
(E-mail, murase%nttlab.ntt.jp@relay.cs.net)

Abstract:
The appearance of a pattern behind a transparent, moving object is distorted by refraction at the moving object's surface. This paper describes an algorithm for reconstructing the surface shape of a non-rigid transparent object, such as water, from the apparent motion of the observed pattern. This algorithm is based on the optical and statistical analysis of the distortions. It consists of the following parts: (1) extraction of optical flow, (2) averaging of each point trajectory obtained from the optical flow sequence, (3) calculation of the surface normal using optical characteristics, and (4) reconstruction of the surface. The algorithm is applied to synthetic and real images to demonstrate its performance.

1. Introduction

One of the primary tasks of a computer vision system [1] is to capture three-dimensional information, such as surface orientation, from two-dimensional images. If some cues are known about the scene, such as stereopsis, shading, contour, texture and motion, two-dimensional images may provide 3-D information (e.g., [1-12]) about the surface. This paper proposes a method for reconstructing the surface shape of an undulating transparent object, such as a water surface, using the cues of refraction and motion. A surface normal is made from the moving, apparent distortion, of patterns viewed through the object. This approach is similar to the method called shape-from-motion. The previous work in shape-from-motion involves reconstructing a 3D structure from movement of several points on the object, based on an assumption of rigidity[7-11]. The method described in this paper does not use the rigidity assumption but uses statistical motion features of the object and optical laws. In addition, the method uses points on refracted images rather than points on the object.

In this method, the objects should have the following two characteristics: (1) they should be transparent with a refraction index not equal to unity, (2) their surface shape should be deformed around an average surface whose shape is known. To clarify the essence of the idea, an example of a water surface with waves[13-14], like the surface of a pool, is used. Because of refraction at the water surface, the observed pattern of the objects under water with waves appears to be moving. Note that human beings can perceive the surface shape from the observed moving pattern. In this case the above two characteristics correspond to the following, (1) water is transparent and has a refraction index of 1.33, (2) the average surface is usually a plane whose surface normal is vertical, and a wave can be regarded as a deformation from the average shape. The goal here is to reconstruct the shape of the water surface from deformed images observed through the waving water, and in so doing, recover the original pattern under the water. The original pattern under the water is assumed to be unknown.

This method has two main original ideas. First, it can be considered as the inverse operation of the ray-tracing approach. The ray-tracing technique is a common method in the field of computer graphics. It is based on an optical law such as a refraction law (Snell's law) or a reflection law, and is used to synthesize the images from the given model (shapes of the objects). For example, P.Y.Ts'o[15] synthesized ocean wave images using physical characteristics and ray-tracing techniques.

Second, this method uses the idea that the pattern observed through the undulating surface is deformed around the pattern observed through the average surface. In the simplest case of the top-view and orthographic projection, the average coordinate of the trajectory of a certain point on the distorted pattern corresponds to the the point observed through the static flat water surface. This means that the average position of the point becomes the position observed without water.

2. Object

First, a brief example is presented. Assume that an image (Fig. 1(a)) is put at the bottom of a pool with waving water, where the wave is a sine wave moving in the horizontal direction. The pattern seen through the water would be distorted and would look, for example, like Fig.1(b) in the case of a smooth wave and like Fig.1(c) in the case of a steep wave. The purpose of our study is first to reconstruct the wave shape from a time series of distorted patterns and finally to detect the original pattern of Fig.1(a).

2.1. Conditions

To make the problem tractable, the following three conditions were introduced.

(1) The amplitude of the wave is small enough that elimination or separation of the pattern does not occur in the observed

(a) (b) (c)

Fig.1. (a) Original pattern,
(b) an observed image in the case of a smooth wave,
(c) an observed image in the case of a steep wave.

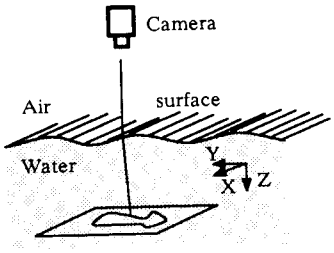

Fig.2. Observation system

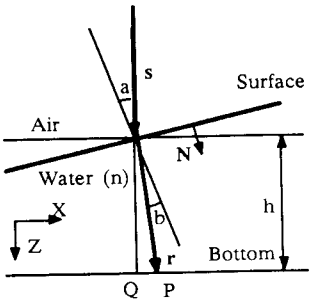

Fig.3. Optical law of refraction

images (i.e. Fig.1(b)). Figure 1(c) is an example containing elimination and separation, so it is beyond the scope of our research at present.

(2) Orthographic projection.

(3) All points in the image are in focus (pin hole camera).

2.2. Assumptions

The following three assumptions are made in the method. Note that the assumption of periodic movement or periodic shape of the surface is not used.

(1) The average slant of the wave surface over a long time period is zero.

(2) Water is transparent and refractive.

(3) The pattern in the water is static.

2.3. Known Parameter

The following parameters, if known, allow the precise surface shape to be determined. If these are not known, the surface shape may still be determined within a scale factor.

(1) The distance, h, between the water surface and the bottom of the tank.

(2) The refraction index, n, of water.

3. Basic Theory

The only physical assumptions introduced are that water is a transparent substance with a certain refraction index, and that the average slant of waves in a long time-scale converges to zero. When a stone falls into water or the wind blows, a complex pattern is generated on the water surface. These wave patterns are generally complicated due to the governing fluid dynamics. To keep the application areas as broad as possible, additional assumptions regarding dynamic behavior have not been introduced, and in fact such assumptions are not necessary for the problem of section 2.

3.1. Optical Characteristics

The refraction law generally known as Snell's law, is effective for this problem among several optical characteristics such as refraction, reflection, polarization and color dispersion. The principle theory of the method based on the refraction law is illustrated in Fig.3. Assuming orthographic projection, a small area of the image, which is located at P, is observed to be at Q due to refraction at the water surface. Here Snell's law is expressed by the vector equation

$$n\,\vec{r} - \vec{s} = \vec{N}\,(\,n\cos b - \cos a\,)\,, \tag{1}$$

where n is the refraction index, $\vec{N}$ is the surface normal, $\vec{s}$ is a unit vector of the ray in air, $\vec{r}$ is a unit vector of the ray in water, a is the incident angle, and b is the refractive angle. The deformation vector $\vec{D}$ is given by

$$\vec{D} = \frac{h\,\vec{r}}{(\,\vec{z}\,,\,\vec{r}\,)} - h\,\vec{z}\,. \tag{2}$$

In order to solve these simultaneous vector equations for $\vec{D}$ using (1) and (2), the unknown vector variables are replaced by the following variables.

$$\vec{s} = (\,0,\,0,\,1\,) \tag{3}$$

$$\vec{r} = (r_x\,,\,r_y\,,\,r_z\,) \tag{4}$$

$$here,\quad r_x^2 + r_y^2 + r_z^2 = 1 \tag{5}$$

$$\vec{z} = (\,0,\,0,\,1\,) \tag{6}$$

$$N = \frac{(\,p\,,\,q\,,\,1\,)}{k} \tag{7}$$

$$here,\quad k^2 = p^2 + q^2 + 1 \tag{8}$$

$$\vec{D} = (\,d_x\,,\,d_y\,,\,0\,) \tag{9}$$

Here, vector s can be expressed by (3) because of the assumption of orthographic projection. If we rearrange the above equations, d_x and d_y are given by the following equations.

$$d_x = h\,p\,\frac{\sqrt{n^2 - k^2 + 1} - 1}{\sqrt{n^2 - k^2 + 1} - 1 + k^2} \tag{10}$$

$$d_y = h\,q\,\frac{\sqrt{n^2 - k^2 + 1} - 1}{\sqrt{n^2 - k^2 + 1} - 1 + k^2} \tag{11}$$

The vector (p, q) is gradient of the surface. k is a weight for normalizing the surface normal, $\vec{N}$. Provided that $k \approx 1$, the water surface is quiet enough. The equations under these conditions are approximated by

$$d_x = h\,p\,(1 - \frac{1}{n}) \tag{12}$$

$$d_y = h\,q\,(1 - \frac{1}{n}). \tag{13}$$

For the case of orthographic projection, the displacement, $\vec{D}$, shows the difference between the position of a point on the observed image and the position on the underwater image. In other words, the observed image is distorted because of the displacement, $\vec{D}$.

3.2. Statistical Characteristics

It is an appropriate assumption for water that the average surface normal converges to a vertical line as time passes. Typical wave patterns, for example, appearing on a pool surface can be assumed as superposed patterns of sine waves with different wave lengths and wave speeds. Such wave patterns can be expressed by

$$f(x, y, t) = \sum_i a_i \sin(u_i\,x + v_i\,y - w_i\,t)\,, \tag{14}$$

where (u_i, v_i) is the wave number, w_i is the angular frequency, a_i is the amplitude, and $f(x, y, t)$ is the wave height at position (x, y) and time t. The surface normal $(p, q, -1)/k$ at a certain point (x, y) can be obtained from the following equations for p and q,

$$p = -\frac{\partial f}{\partial x} = -\sum_i a_i \cos(u_i\,x + v_i\,y - w_i\,t) \tag{15}$$

$$q = -\frac{\partial f}{\partial x} = -\sum_i a_i \cos(u_i\,x + v_i\,y - w_i\,t)\,. \tag{16}$$

Thus, average value p' and q', regarding time, of p and q are expressed by

$$p' = \frac{1}{T}\int_0^T p\ dt \tag{17}$$

$$q' = \frac{1}{T}\int_0^T q\ dt. \tag{18}$$

Assuming that wi is never equal to zero, p' and q' will converge to zero as T increases. In special situations, such as a continuous spout of water, w_i in (17) (18) can be zero. In such a case p' and q' would not converge to zero, and the accuracy of the shape reconstruction would be reduced.

4. Reconstruction Algorithm

4.1. Outline of the Algorithm

The purpose of the algorithm is to reconstruct the shape of the water surface and to detect the image under the water. It consists of the following 4 steps: (1) extraction of optical flow, (2) averaging the trajectory of each point in the observed images, (3) calculation of the surface slant using optical characteristics, and (4) shape reconstruction. Each step is detailed in the following experiments using synthesized images.

4.2. Synthesized Image

Synthesized images were used to demonstrate the method. The images were synthesized by the ray-tracing method under the following conditions. This is one of the examples. Let the pattern of Fig.1(a) 50cm × 50cm be an underwater pattern. The depth of water is 50cm. The sine wave moves diagonally from the upper-left of the image to the lower-right along the water surface. The maximum slant is 5.0 degree, the wave length is 16.6cm (3 cycles in 50cm), the wave speed is 2.1cm/frame (8 frames = 1 cycle). Fig.4 shows a sequence of the wave surface shape for synthesizing the images (frames 1,2,3). The viewing point is high enough that the projection can be approximated as orthographic. Figure 5 shows dynamic images (frames 1,2,3,4) obtained by this method. In this case the surface shape is periodic. However, the algorithm can also reconstruct a water surface which is not clearly periodic.

4.3. Extraction of Optical flow

Optical flow expresses the movement of a certain point in the image from one time to another. For computation, the gradient method (e.g.,[3]), the correlation method (e.g.,[16]) and the token tracking method are well known. Various methods were used for various patterns. A correlation method, a simple algorithm, is applied to these synthesized images, since correlative features are preserved throughout the series of images at different frames in the experiment. The correlation value $d(i', j')$ between the rectangle (with side length a) nearest to (i,j) in the t-th frame and the rectangle nearest to $(i+i', j+j')$ in the $(t+1)$-th frame is determined from the following equation.

$$d(i', j') = \sum_{-\frac{a}{2} < k,l < -\frac{a}{2}} |f(i+i'+k, j+j'+l, t+1) - f(i+k, j+l, t)| \tag{19}$$

The vector (i', j') that makes $d(i', j')$ smallest is considered as optical flow of (i,j) at time t. Figure 6 shows an example of an optical flow map extracted by this method.

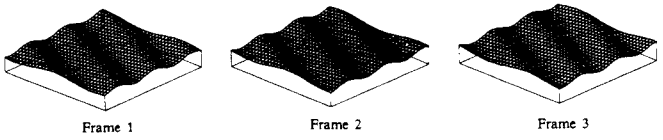

Fig.4. Sequence of wave surface shape for synthesizing images.

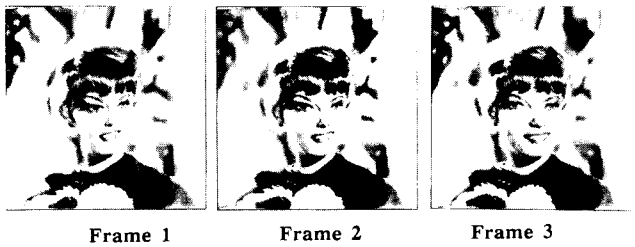

Fig.5. Synthesized image sequence distorted by refraction.

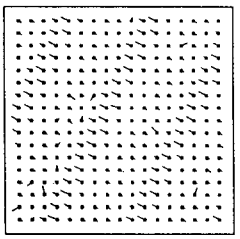

Fig.6. Extraction of optical flow.

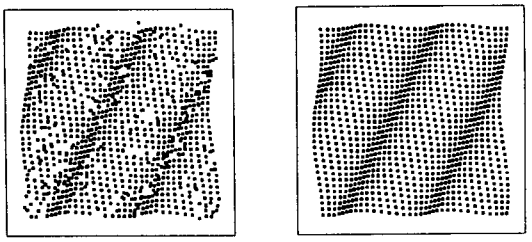

Fig.7. (a) Calculated center-of-trajectory.
(b) Theoretically correct center-of-trajectory.

4.4. Center of the Trajectory (COT)

In this section, we define "Center of trajectory (COT)", which is the main point of the algorithm, and we suggest a method using COT to recover the underwater pattern, the second purpose of this paper. Consider a certain point P_t in the observed image at time t. The average coordinate (x,y) of the trajectory of P_t, which is derived from optical flow, is defined as the COT of point P_t. This corresponds to the position of P_t observed through water whose surface has an average shape with time. The average surface normal generally converges to zero (horizontal flat surface).

Figure 7(a) shows the obtained COTs of grid points at frame 1 of the above synthesized images, while Fig.7.(b) shows the theoretically correct COTs.

The underwater pattern can be recovered by using the COT. The image obtained by mapping all the pixels of an observed image at a certain time on their corresponding COT will be the image as seen through a flat water surface, if orthographic projection can be assumed. The recovered image of a synthesized sequence is shown

Fig.8. Recovered underwater pattern.

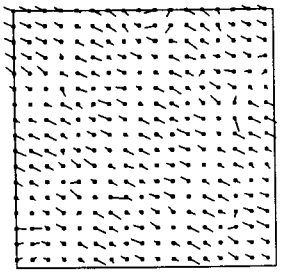

Fig.9. Vector field showing surface normal.

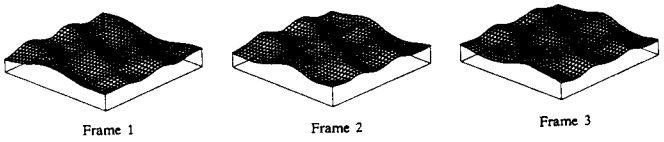

Fig.10. Shape reconstruction of water surface.

in Fig.8. The remaining distortion in this image is only due to incorrect extraction of optical flow.

4.5. Gradient Vector Field

If the points (x_t, y_t) and COT (x', y') of a point are known, the direction of vector normal to the water surface (p_t, q_t) at (x_t, y_t) can be calculated from equations (10) and (11). The gradient at (x_t, y_t) is thus obtained from

$$P_t = \frac{(x_t - x')\, n}{h\,(n-1)} \tag{20}$$

$$q_t = \frac{(y_t - y')\, n}{h\,(n-1)}. \tag{21}$$

It requires a large amount of computation to extract the optical flow vector for all points in the image. Therefore optical flows, COTs and gradient vectors are first computed corresponding to the selected points, and the gradient vector of each pixel is calculated by averaging the gradient vector of four points adjacent to the pixel. Gradient vectors of different points in image frame 1 are shown in Fig.9. Here the vectors are thinned out for clarity.

4.6. Shape Reconstruction

When the gradient vectors of each point are given, the surface shape can be reconstructed by integrating them over the X and Y dimensions. The reconstructed surface shapes for frames 1,2 and 3 are illustrated in Fig.10. The Z direction has been expanded for clarity.

5. Experiment with Real Images

5.1. Collecting Data

A TV camera was used to collect data. First, 2mm × 2mm black dots were randomly scattered on a blank sheet of paper by a computer. The paper was fixed on the bottom of a water tank with water depth h=25cm. This situation is similar to one where black stones are scattered randomly in a pond. Waves were then induced by stirring the water. A video camera was placed 1 m from the water surface. The camera produced an image of 512 times 480 pixels and the focal length was fixed so that one pixel on the sampled image corresponded to 0.5mm × 0.5mm on the underwater pattern. 90 frames of dynamic images were collected at a rate of 30 frames per second for 3 seconds. Figure 11(a) and (b) are examples of original images (frame 1,2).

5.2. Extraction of Optical Flow

In case of these images, black pixel clusters are useful tokens for optical flow extraction. The process based on the token tracking method is as follows. First, the shading corrected images are binarized and then parts which have more than a specified number of black pixels are picked up. The optical flow is extracted by tracing these parts. The optical flow obtained by this method was manually checked and found to contain no errors. Figure 12 is an optical flow map extracted from frame 6.

5.3. Determination of COT

Figure 13 shows the trace of a certain point on the image for 0.66 sec. (20 frames) using optical flow. From this result, it can been seen that the locus of a point moves around a certain point of the real image. The COT can be calculated in the same way as explained in 4.4. The result shown in Fig.14 shows the pattern of black points that actually lies at the bottom of the water.

5.4. Computation of Gradient Vector Field

The gradient vector field was calculated using the positional relation between a point and its COT, as explained in 4.5. As an example, the gradient vector field for frame 6 is shown in Fig.15.

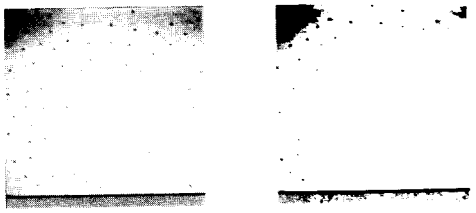

Fig.11. Observed image sequence.

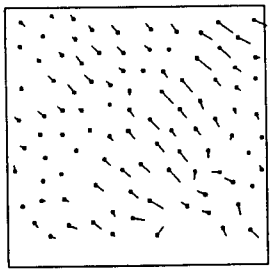

Fig.12. Optical flow.

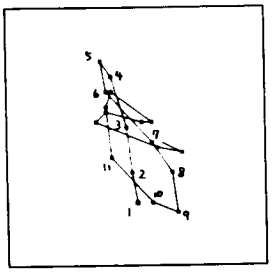

Fig.13. Tracking of a point on the pattern.

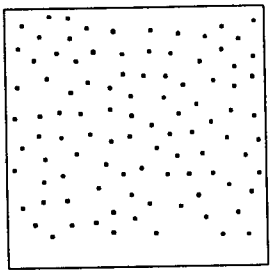

Fig.14. Center-of-trajectory.

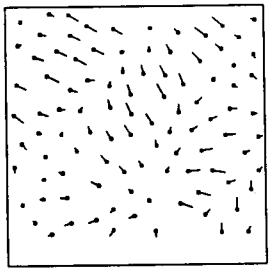

Fig.15. Vector field showing surface normal.

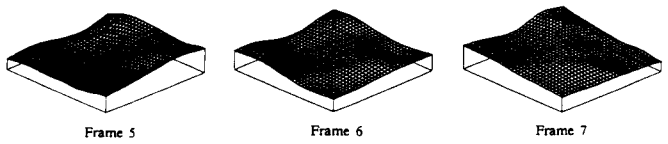

Fig.16. Sequence of reconstructed surface shape.

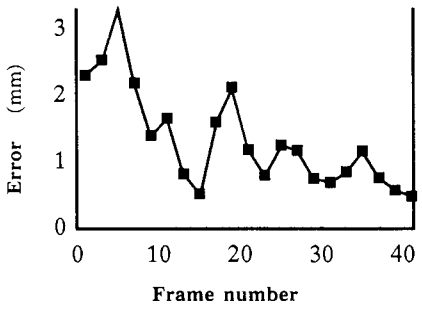

Fig.17. Estimated errors of the surface shape reconstruction.

5.5. Shape Reconstruction

The shape of the wave was reconstructed from the gradient vector field using the method of 4.6. Figure 16 shows the reconstructed wave shapes for frames 5 through 7. Thus, it has been demonstrated that the wave shape can also be reconstructed for a real image sequence. Therefore, the algorithm presented here is useful for shape reconstruction in a passive observation system.

6. Shape Reconstruction Errors

The errors are mainly due to (1) failure in extracting optical flow, and (2) errors in calculating COT. To improve (1) a better calculation algorithm needs to be developed. Here we concentrate on the evaluation of errors (2). Errors can be minimized by prolonging the observation time. Results are presented below for experiments with real images (the same data in section 5). We assume that the available information for reconstruction of the surface is taken from the first frame (t=1) through the present frame. The error in the t-th frame, S_t, is defined as

$$S_t = \frac{1}{M} \sum_{i,j} | Z_{tij} - Z'_{tij} |, \qquad (22)$$

where M is the total number of combinations of i and j, Z_{tij} is the Z value at coordinate (i,j) at the t-th frame of the reconstructed shape by our method, and Z'_{tij} is the Z value of the standard surface shape. The standard surface shape is calculated using the actual underwater dot pattern. As shown in Fig.17, the error decreases with repeated oscillation. The phenomenon of increasing error during the first few frames can be explained as follows. The position of points tends to oscillate, so some points move in a different direction from the COT. Where such points predominate, the error first increases, but then decreases after a long period.

7. Conclusions

The movement of a transparent object past a pattern, distorts the observed image of the pattern. This paper dealt with the surface shape reconstruction of the transparent object from a sequence of distorted 2D images. It can be thought of as an inverse operation of the ray-tracing technique, which is well known in Computer Graphics. The proposed method can deal with objects having the following 2 characteristics: (1) they are transparent and have a non-unity refraction index, (2) their surface shapes are deformed around an average surface whose shape, in relation to time, is known. Its originality lies in dealing with a nonrigid transparent object and in using cues of refraction (optical law, the above characteristic (1)) and motion (statistical characteristics of the object, the above characteristic (2)). A example of a water surface with waves was used. The shape of an undulating transparent object and the original pattern behind the object can be recovered simultaneously, from the time series of the distorted images. The estimation error decreases as time passes. The method has been demonstrated in experiments using both synthesized images and real images. Areas for future research include improving the accuracy of optical flow extraction and applying the method to various kinds of real images.

Acknowledgment I would like to thank Dr. S. Naito and Dr. J. Roseborough for valuable advice in preparing the paper. I am also grateful to Mrs. Y. Murase for her linguistic collaboration.

[References]

[1] H.Ballard and C.M.Brown, Computer Vision, Prentice-Hall, 1982.
[2] W.E.L.Grimson, From images to surfaces, A Computational study of the human early visual system, MIT Press, Cambridge, MA, 1981.
[3] B.K.Horn, Robot Vision, MIT Press, Cambridge, MA, 1986.
[4] A.P.Witkin, "Recovering surface shape and orientation from texture", Artif. Intell., vol.17, PP.17-45, 1981.
[5] J.Aloimonos, "Shape from texture", Biol. Cybern., vol.58, pp.345-360, 1988.
[6] H.C.Longuet-Higgins, "The visual ambiguity of a moving plane", Proc. Roy. Soc. Lond,. Ser.B, vol.223, pp.165-175, 1984.
[7] M.Subbaro "Interpretation of image flow: Rigid curved surface in motion", Inter. Journal of Computer Vision, 2, pp.77-96, 1988.
[8] K.Kanatani "Structure and motion from optical flow under perspective projection", Comp. Vision Graphics Image Processing, vol.38, pp.122-146, 1987.
[9] H-H.Nagel, "Representation of moving rigid objects based on visual observations", IEEE Computer, 14, pp.29-39, 1981.
[10] O.D.Faugeras, F.Lustman and G.Toscani, "Motion and structure from motion from point and line matches" in Proc. Int. Conf. Computer Vision, pp.25-34, 1987.
[11] S.Ullman, The interpretation of visual motion, MIT Press, Cambridge, MA, 1979.
[12] D.Marr, Vision, W.H.Freeman and Company, New York, 1982.
[13] B.L.Gotwols and G.B.Irani, "Charge-coupled device camera system for remotely measuring the dynamics of ocean waves", Applied optics, vol.21, No.5, March, pp851-860, 1982
[14] B.Jahne, "Image sequence analysis of complex physical objects: Nonlinear small water surface waves", in Proc. Int. Conf. Computer Vision, pp191-200, 1987.
[15] P.Y.Ts'o and B.A.Brasky, "Modeling and rendering waves: wave-tracing using beta-splines and reflective and refractive texture mapping", ACM Trans. on Graphics, vol.6, No.3, July, pp.191-214, 1987.
[16] J.M.Prager and M.A.Arbib, "Computing the optic flow: the MATCH algorithm and prediction", Comp. Vision Graphics Image Processing, vol.24, pp.213-237, 1984.