

Range Data Reconstruction Using Fourier Slice Theorem*

Shih-Schön Lin and Chiou-Shann Fuh

Department of Information Engineering and Computer Science

National Taiwan University

E-mail: r3506011@image.csie.ntu.edu.tw, fuh@csie.ntu.edu.tw

Abstract

This paper proposes a new approach to resolve the ambiguity problem in multistripping laser triangulation systems.

Our approach is based on the Fourier slice theorem that would be briefly described below. This theorem also forms the basis of X-ray CT (Computed Tomography) reconstruction.

1. Introduction

To recover the third dimension out of 2D pictures, one series of techniques the precursors in this field devised is:

- Structured lighting to resolve line of sight [5].
- Multistripped laser pattern to increase scanning area [3].
- Optical shutter pattern to resolve ambiguity of multistripped laser pattern [2].

The principle of structured light ranging systems is triangulation, see Fig. 1. Stereo pair ranging utilizes triangulation as well, but it is much more difficult to match corresponding points out of the stereo image pairs. Structured light method, on the other hand, could easily pinpoint the position of a distinctive light stripe (Fig. 2) and resolve the depth. However, as

*This research work was supported by National Science Council of Taiwan, ROC, under Grants NSC 84-2212-E-002-046 and NSC 85-2212-E-002-077, by Cho-Chang Tsung Foundation of Education under Grant 84-S-26, by Institute of Information Industry under Grant III 84-TC-0422, and by Mechanical Industry Research Laboratories, Industrial Technology Research Institute under Grant MIRL 0988.

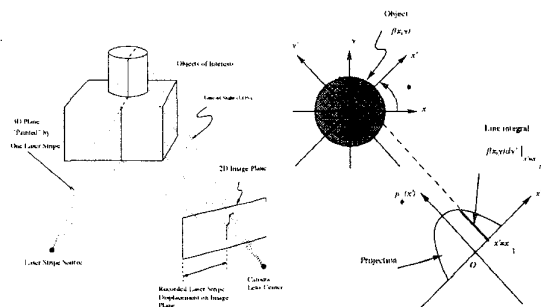


Figure 1. (a) The role of laser in triangulation. (b) Radon transform or projection data of 2D surface function $f(x, y)$ at a given view angle ϕ .

we increase the number of stripes we see in one image, another ambiguity problem arises. We would discuss this problem in greater detail in the next section, as this is the major problem we are going to solve.

Fourier slice theorem is the backbone of X-ray Computed Tomography (CT) [1]. This theorem states that if the projection function (Radon transform) of a 2D surface $f(x, y)$ is known at every angle, one can uniquely recover individual value $f(a, b)$ for any point (a, b) . We would show that the problem we discuss can be reduced to the recovery of a conceptual $f(x, y)$, and we devise a new approach to solve our problem with Fourier slice theorem.

2. Ambiguity problem and optical shutter solution

Suppose we are inspecting industrial parts on a conveyor belt. Before the parts are fed in, the camera sees approximately equally spaced laser stripes on the

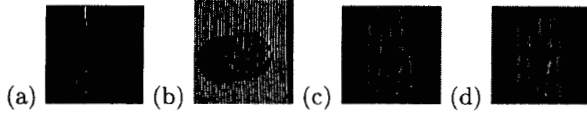


Figure 2. (a) Single laser stripe projection. (b) Typical image of multi-stripe laser triangulation system. (c) (d) Two views of a toy plane.

flat conveyor belt. We call the image positions of each stripe the starting positions of the stripes when there are no objects on the conveyor belt. Once the objects of interests come into inspection area, our camera would see the positions of laser stripes move. We call the image positions after the objects of interests enter the inspection area the ending positions of each stripe. Such displacements in our 2D raw data are related to the 3D depth information we want.

To calculate displacements we need both the starting positions and ending positions. Unfortunately, raw data give only the ending positions. When we use only one single laser stripe, we get displacements easily because we control one and only one stripe so we know where each starting position is before the object comes into sight. Once we project more than one stripe at the same time, we lose track of which starting point corresponds to which ending point. This is the so called multistriping ambiguity problem. Note that we still know all the values of the starting and ending positions in multistriping systems, we lose only their correspondences. This is the major clue for our method to recover the lost information.

A mechanical or electrical solution to the multistriping ambiguity problem is adding optical shutters in front of each of the stripes we project. Optical shutters allow more control over the light stripes. By turning on one stripe at a time, the ambiguity disappears. This time we do not have to scan the stripe or move the object, which reduces the errors introduced in the scanning process. If we number the stripes and “switch” on and off the bits of each number, then we can reduce the number of images taken in one “scan” to the logarithm of the number of stripes provided, see [2].

Unfortunately, one may not always have easy access to the specific control electronics and optical components for shuttering. Moreover, if we can do the job without adding mechanical or electronic gadgets we save money and increase system reliability at the same time. Should the sophisticated controls needed by existing algorithms break down at crucial moments, those algorithms cease to function as well. Our proposed solution fills in this gap.

3. Mathematical background

3.1. Perspective projection

Basically each point in an image represents one line of sight (LOS), see Fig. 1 (a). When the center of perspective projection is distant from the image plane, perspective projection can be approximated by orthogonal projection.

3.2. Radon and Fourier transform

Radon transform operator for a given function f is defined as

$$\begin{aligned} p_\phi(x') &\equiv \mathcal{R}[f(x, y)] \\ &= \int \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \phi + y \sin \phi - x') dx dy \\ &= \int_{-\infty}^{\infty} f(x' \cos \phi - y' \sin \phi, x' \sin \phi + y' \cos \phi) dy' \end{aligned}$$

where

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

2D Fourier transform of an image function $f(x, y)$ is defined as

$$F(\omega_x, \omega_y) = \int \int_{-\infty}^{\infty} f(x, y) \exp\{-i(\omega_x x + \omega_y y)\} dx dy$$

4. Proposed new solution

For illustration let us consider the case that we project $2p + 1$ stripes which form straight lines parallel to the y axis of the image when there is nothing but the calibration board on our inspection table. Here we assume the camera is looking right on top of the inspection table, and orthogonal projection can be used to approximate perspective projection. Also we assume the image taken can resolve $2n + 1$ data points. For convenience of discussion and ease of implementation, we assume the undistorted laser stripes are seen as several columns of bright pixels or rows of bright pixels. This is easily achievable in our system, see Fig. 2 (b). Then we can construct a conceptual discrete 2D domain with x ranging from $-p$ to p and y ranging from $-n$ to n , see Fig. 3. We use the function $f(x, y)$ to calculate the 3D depth:

$$f(x, y) = l_1(x, y) - l_0(x, y)$$

where $l_0(x, y)$ is the position of x th laser stripe at y th row on image when there is no object on the inspection

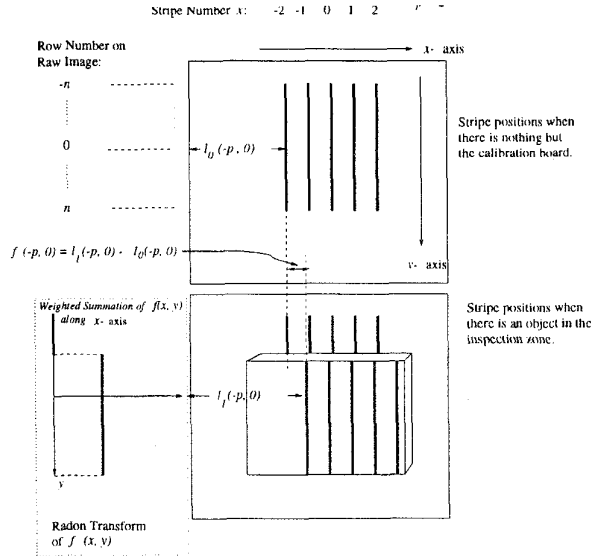


Figure 3. The physical meaning of conceptual function $f(x, y)$

table, whereas $l_1(x, y)$ is the position of x th laser stripe at y th row on image when there is object on the inspection table. We know exactly the function value $l_0(x, y)$ for every (x, y) in our domain because the spatial ordering of laser stripes in calibration image is under our control, but we usually can not predict the ordering when the inspected objects come into play. In short, suppose we still get $2p+1$ strips and their positions, we are still unsure which value belongs to which $l_1(x, y)$.

Now let us take a look at its discrete Radon transform (See Fig. 1 (b)) along the y axis of our domain:

$$\begin{aligned}
 p_{\frac{\pi}{2}}(x') &\equiv \mathcal{R}[f(x, y)] \\
 &\Rightarrow \sum_x \sum_y f(x, y) \delta(x \cos \frac{\pi}{2} + y \sin \frac{\pi}{2} - x') \Delta x \Delta y \\
 &= \sum_{y'} f(-y', x') \Delta y' \\
 &= \sum_{\xi} f(\xi, x') \Delta \xi \\
 &= \sum_{\xi} l_1(\xi, x') \Delta \xi - \sum_{\xi} l_0(\xi, x') \Delta \xi
 \end{aligned}$$

Both the last two summations are calculable regardless of whether we can match the values to their respective $l_1(x, y)$ or not.

Assume that the object surface to be inspected is sufficiently smooth so that we can approximate its Radon transform of our conceptual domain at many

other angles by taking pictures as we change the relative angles between the inspection imaging system and the object, then we can proceed with the application of Fourier slice theorem to resolve $(2n+1) \times (2p+1)$ pairs of correspondences altogether.

Consider the 1D Fourier transform of Radon transform

$$\begin{aligned}
 P_{\phi}(\omega) &\equiv \mathcal{F}_1[p_{\phi}(x')] \\
 &= \int_{-\infty}^{\infty} p_{\phi}(x') \exp(-i\omega x') dx' \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp[-i\omega(x \cos \phi + y \sin \phi)] dx dy \\
 &= F(\omega \cos \phi, \omega \sin \phi) \\
 &= F(\omega_x, \omega_y)|_{\phi} \\
 &= F(\omega, \phi)
 \end{aligned}$$

Note that \mathcal{F}_1 stands for one dimensional Fourier transform. The last form is the 2D Fourier transform of the conceptual function $f(x, y)$ presented in Fourier domain polar coordinates. This means that the 1D Fourier transform of the projection data (Radon transform) at a given view angle is equivalent to the radial data passing through the origin at a given angle ϕ in the 2D Fourier transform domain data.

So if we get sufficient sampling in many view angles or we know in advance that the object possesses certain forms of symmetry, then we could in principle fill out the 2D Fourier transform of our intended domain and inverse transform to reconstruct our desired function.

So far, the theorem is derived purely from continuous domain, but all we get in the real world application is sampled discrete data. So techniques like Filtered Backprojection (FB) Algorithm [1] are developed to handle these "real world imperfections".

Consider the inverse 2D Fourier transform in frequency domain polar coordinates (ω, ϕ) and space domain polar coordinates (r, θ) :

$$\begin{aligned}
 \hat{f}(r, \theta) &= \int_0^{2\pi} \int_0^{\infty} F(\omega, \phi) \exp[i\omega(x \cos \phi + y \sin \phi)] |\omega| d\omega d\phi \\
 &= \int_0^{\pi} \int_{-\infty}^{\infty} |\omega| P_{\phi}(\omega) \exp[i\omega(x \cos \phi + y \sin \phi)] d\omega d\phi \\
 &= \int_0^{\pi} p_{\phi}^*(x') d\phi
 \end{aligned}$$

where

$$p_{\phi}^*(x') = \int_{-\infty}^{\infty} |\omega| P_{\phi}(\omega) \exp(i\omega x') d\omega$$

$$\begin{aligned}
&= \mathcal{F}_1^{-1}[\omega | P_\phi(\omega)] \\
&= \mathcal{F}_1^{-1}[\omega] * p_\phi(x')
\end{aligned}$$

Note here that \mathcal{F}_1^{-1} stands for one dimensional inverse Fourier transform. In actual implementation, we would use discrete Fourier transform (DFT) to obtain $P_\phi(\omega)$ and inverse discrete Fourier transform (iDFT) to obtain $p_\phi^*(x')$ [4]. Another approach is to derive $\mathcal{F}_1^{-1}[\omega]$ analytically and sampling the results in the spatial domain. Note here that either approach we get the values of $p_\phi^*(x')$ only for a fixed number of x' 's. However, when we evaluate $\hat{f}(r, \theta)$, we need for each ϕ the value of $p_\phi^*(x')$ at $x' = x \cos \phi + y \sin \phi$. In such case, we can use simple linear interpolation to obtain acceptable results. Also, in this manner we could generate arbitrarily dense values for $\hat{f}(r, \theta)$. This is a favorable characteristic in multimedia applications, as rough scanning is much faster than dense scanning, but could still yield dense range images most users want.

To recover the original heights, certain rescaling is needed. If only the matching relation between the starting positions and ending positions is requested, the rescaling could be omitted, as the relative strength between $f(x, y)$ is sufficient for our purpose.

Another major problem involved is that we usually do not get perfect projection at every scan line. Along one scan line, a laser stripe may be missing due to shadows (e.g. see Fig. 2 (b)) or because the line displacement exceeds image boundary (e.g. see Fig. 2 (a) around the center). In this case we simply treat that scan line as more sparsely sampled with respect to other scan lines (see Fig. 4 (a)). That is, since we are performing integration on computer, we actually use one value in an integration interval to represent all point values in the interval, so when we get less data points we simply extend the length of interval for each point. As we do not know which starting position should be dropped, we drop values in multiples of their average. In this way we greatly reduce the impact of missing lines. Our experiments demonstrate that our method can still recover most shapes when there are considerable amount of missing lines.

5. Experiments and simulations

The lenses we use include Micro-Nikkor 200mm f/4, 55mm f/2.8, and Kiron 105mm f2.8. The laser light source is Lasiris model ULF-533L which can produce 33 parallel lines simultaneously. We use Sun Sparc station 10 for reconstruction computation that takes about 20 seconds when we reconstruct 128 projections. When reconstructing Fig. 4 (b) from Fig. 2 (b) the single projection extracted from real image can serve

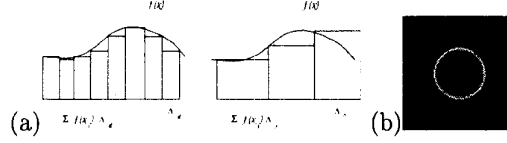


Figure 4. (a) Densely sampled digital integration (left), and sparsely sampled digital integration (right). (b) Recovered range data for Fig. 2 (b) (rms = 10.7, dre = 0.07 %). We used only one picture to reconstruct this range image.

as each of the 128 projections, thanks to the circular symmetry property of this object. The centroid of a cluster of stripe pixels is treated as the position of a laser stripe. To avoid introducing too many factors (e.g. noise, instrumentation faults, ..., and so on) at the same time, we create a series of simulated range images and a striping simulator to demonstrate the response of our method to individual imperfections. Missing lines, asymmetries are added one at a time. Fig. 5 shows the objects to be reconstructed. Fig. 6 shows the original and reconstructed range image of the first model. This model is both circular symmetric and shadow-free. Fig. 7 shows the effects of shadows and the effects of both shadow and asymmetry to our reconstruction algorithm. The (f) is reconstructed using 80 angles. It is clear from the results that the output degrades as shadow increases.

6. Comparison and discussion

The best case for our method is circularly symmetrical objects, in which case only one picture is enough to resolve any ambiguity and the complexity relative to the number of stripes used is constant. Not everything is circularly symmetric, but circular symmetry is an often possessed property among manufactured objects, and these mass produced parts need machine vision inspection most.

Another feature of the proposed method is its ability to generate dense range images. Optical shuttering solution use $\log N$ pictures to resolve the ambiguity of "one" picture. All $\log N$ pictures are taken at exactly the same relative position between the inspection system and the object. This is rather wasteful if we want the complete surface shape of an object. For general objects our method needs many more pictures than $\log N$, but each picture used views the object from a different angle. When our algorithm completes, all ambiguity in each multistripe picture is solved simultane-

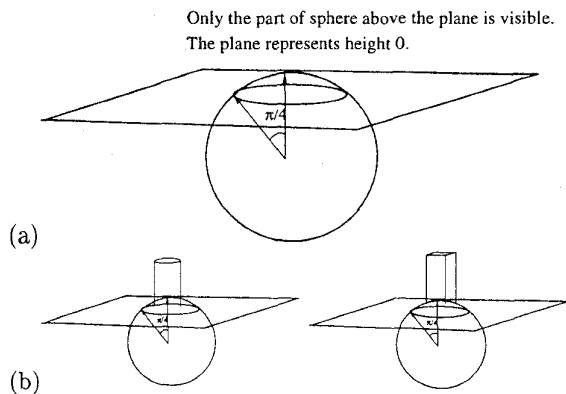


Figure 5. The models used in the simulation. (a) The side view of the first model, representing shadow free 2D circularly symmetric case. (b) Left: adding to (a) a cylinder concentric to the original sphere would produce significant shadow, however, circular symmetry is still valid. Right: adding a square column instead of a cylinder, shadows are introduced and circular symmetry is no longer valid.

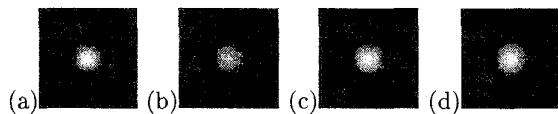


Figure 6. Gamma corrected range images: (a) Original partial sphere model. (b) Reconstructed image of (a) using 128 projection angles (rms = 0.4). (c) Reconstructed image of (a) using only one projection (rms = 0.2), here the circular symmetry characteristic is exploited. Furthermore, only one half of the points are used to calculate the projection, but using our normalization method the result is almost identical as that of (d). (d) Like (c), we use only one projection, but this time all available points are used to calculate the projection (rms = 0.2). (rms: root mean square error, dre: distance relative error)

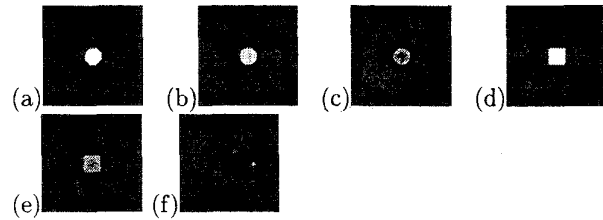


Figure 7. Gamma corrected range images: (a) Original range image of the model with one cylinder above the sphere. (b) Reconstructed range image using only one projection (rms = 1.2). (c) Reconstructed image using 128 projections (rms = 2.1). (d) Original range image of the model with one square pillar above the sphere. (e) Reconstructed range image using 128 projections (rms = 2.5). (f) Reconstruction from Fig. 2 (c)(d) (rms = 15.3, 1.5mm, dre = 0.1%).

ously, and we get much more range data points than the optical shuttering solution.

7. Conclusion

When complete profile of a side of object is needed, the proposed resolution is better than most existing ones. Furthermore, when circular symmetry is present, our method requires the fewest number of pictures—one.

References

- [1] Z. H. Cho, J. P. Jones, and M. Singh. *Foundations of Medical Imaging*. Wiley, New York, 1993.
- [2] R. Jain, R. Kasturi, and B. G. Schunck. *Machine Vision*, pages 300–304. McGraw-Hill, New York, 1995.
- [3] J. A. Jalkio, R. C. Kim, and S. K. Case. Three dimensional inspection using multistripe structured light. *Optical Engineering*, 24:966–974, 1985.
- [4] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*, pages 498–414. Cambridge University Press, Cambridge, UK, 1988.
- [5] P. M. Will and K. S. Pennington. Grid coding: A novel technique for image processing. *Proceedings of the IEEE*, 60:669–680, 1972.