Development of a 3D Vision Range Sensor using Equiphase Light-Section Method

KUMAGAI Masaaki
Department of Mechanical Engineering
Tohoku Gakuin University
Tagajo, Japan

Abstract—Many of autonomous robots or industrial robots require three-dimensional environmental sensors. Stereo visions or laser range finders are often used to detect distance to obstacles or targets. A new type of three-dimensional measurement system is proposed in this paper. This method is a kind of active stereo vision based on light-section method and phase-shift method. A special light source that consists of a rotating cylinder and a lamp are used. It projects planes of lights whose intensities vary in a sinusoidal function of time. Phase of each plane is different from those of other plane, and we can distinguish which plane sectioned the target from the phase. The images obtained from video camera are processed to be phase of each pixel that indicates distance to measuring targets. The method needs only simple calculation for the detection of phase. The basic principle and expanded methods for practical use are also proposed in this paper. The experimental results using trial system are also given, which show effectiveness of the method.

I. INTRODUCTION

A stereo vision became one of most famous environmental sensors for mobile robots such as wheeled robots or walking robots. Generally, the stereo visions for these robots are passive stereo visions, i.e. they have two video cameras (their optical axes are often parallel to each other) without lighting system, and a pair of images is processed with software or hardware to obtain range image. The author and colleagues had developed and used such stereo vision for a biped robot to detect condition of ground surface and obstacles[1], [2]. A similar system had been also applied to a quadruped robot that can step over low-height obstacles[3]. An Autonomous land vehicle was also equipped with stereo vision system[4].

However, the author had been aware of limitation of the stereo vision. The stereo matching calculation, that is main operation of the system, requires rich and unique feature on target. We cannot detect matching point between right and left image if the obtained images are featureless, and mismatching will occur if the images contain repeated patterns such as net fence or well brick-topped road, whereas robots must work in such scenes especially in indoor environment. Moreover, it takes large computation power, i.e. computation time or expensive hardware. From these points of view, the author planned to develop an active 3-D vision system that consists of camera(s) and special light source(s).

Many types of active 3-D vision system have been proposed. They can be classified into the method based on analysis of Moire patterns, passive stereo vision with pattern projection[5], [6], light-section method[7], phase shift method[8], [9], and etc. The light-section methods and phase-shift methods have been improved. Especially in industrial application, phase-shift methods seems to be used widely[9].

In this paper, a 3-D vision ranging system that looks like a hybrid between light-section method and phase-shift method. The basic idea of this method is brought from the light-section method. The light section method projects a slit light and detects the shape of the lighten line. Only one or a few lines can be projected at the same time because of separation of lines. The idea of proposed method is substitution of varying intensity lights for slit lights. The light source generates many (and continuous) sectioning equiphase-planes whose intensities are decided by sinusoidal function of time. Each of sectioning line can be easily distinguished because each of equiphase planes has different phases. The phase-measuring technique as same as that of phase-shift method is used to detect which equiphase plane crosses each of pixels in observing camera image. The detected phase is converted into distance.
shown.
first prototype of detection system and obtained results are
purpose. The principle of detection is described, and then,
light source, and striped pattern for each of detection
camera (resolution, speed, wavelength of light), type of
rotating striped cylinder. We can easily select appropriate
information to obtain 3-D map.
This vision system consists of an ordinary video camera
and a (or a few) light source such as electric bulb with
rotating striped cylinder. We can easily select appropriate
camera (resolution, speed, wavelength of light), type of
light source, and striped pattern for each of detection
purpose. The principle of detection is described, and then,
first prototype of detection system and obtained results are
shown.

II. PRINCIPLES OF DETECTION

A. Basic idea

Figure 1 shows principle of the light-section method. A
slit light beam (a plane of light) is projected to measuring
target. The lighten line (shown in thick line) will be
observed from camera as curved and/or segmented lines
depending on environment. The 3-D shape of the target can
be recognized by processing shape of the lighten line. The
number of slit light that can be projected simultaneously
is limited to only one or a few because it is not easy to
distinguish segmented lines if there are many lines. One
of current approach to speed up the light-section method
is to speed up detection of one line and to make detection
much frequently[7].

To solve the limitation, the light plane with phase is
introduced in the equiphase method. Many light plane
whose intensities vary in sinusoidal function with different
phases are projected simultaneously as shown in Fig. 2. The
lighten line in light-section method becomes a set of points
whose brightness change with same phase, an equiphase
line. If we process the shape of equiphase line, we obtain
shape of target as same manner as light section method. Of
course, it takes a few to a dozen image fields to project and
detect phase of brightness, which takes time. The advantage
of this method is that the detection process of phase is
linear calculation as described below whereas lighten line
detection in light-section method needs binalization whose
threshold value makes us feel melancholy. Therefore, the
equiphase method is not affected by color of target (of
course, black object is not suitable), and background lights.

We can use much easier approach as shown in Fig. 3.
We assume that many equiphase planes are projected
continuously. Some equiphase plane lightens a point on the
target if the point is not in shade. Different phase is
observed from camera on the point according to distance
from camera to the point. For example, if a object is locate
at (b) in Fig. 3, the phase \( \phi_2 \) on point B will be detected on
center pixel of camera, whereas if the object is at (a) or (c),
the phase \( \phi_3 \) or \( \phi_1 \) is obtained respectively. This means that
we only have to detect the phase of brightness of observed
pixel in camera image to obtain distance information. The
phase can be converted to distance by geometry, by looking
up prepared table, by comparing with reference object, etc.
with system parameters.

The light source that projects continuous equiphase
planes is materialized with a bulb and rotating striped
cylinder as shown in Fig. 4. The transmittance of striped
patter \( \alpha \) varies as a function of central angle \( \psi \) from a
reference point. For example, the pattern described with function

\[
\alpha(\psi) = 0.5 + 0.5 \sin(2\psi)
\]

gives us equiphase-plane whose intensity changes two
cycles per one revolution of the cylinder. Though actual
phase of each plane is decided by direction of plane and
where we put reference of time and angle, there are no two
planes whose phases are same in any area within 180 [deg]
of lighting direction. If the uniqueness of the phase is
guaranteed in measuring area, the phase can be converted
directly into distance.

B. Detection of phase

The detection of the phase from camera images is
described hereafter. As described above, a rotating cylinder
with stripes is used. We assume the light projected to
certain direction has intensity as follows,

\[
I = I_0(a + b \sin(2\pi ft + \phi))
\]

where \( I \) and \( I_0 \) are intensity of a plane and maximum
intensity, \( a \) and \( b \) are coefficient of transmittance of even
and amplitude \( (a + b = 1, a - b \) gives minimum), \( f \)
is time, and \( f \) is frequency that depends on rotational
speed and number of cycles of striped pattern on cylinder,
respectively.

The sequence of images of target under this light is taken
by video camera. The sequence of brightness of a pixel on
image located at coordinate \( (x, y) \) can be obtained as

\[
Y_{xy}[t] = Y_{0,xy}(a + b \sin(2\pi i/n + \phi)) + Y_{B,xy}
\]
where \( Y_{xy}[i] \) is brightness of the point on image field \( i \), \( Y_{0,xy} \) is maximum brightness, and \( Y_{B,xy} \) is brightness brought by background lights, respectively \((x, y, i \) are integer values). The constant \( n \) is field count needed for one cycle of pattern. For practical system, the rotation of the cylinder must be synchronized to video field so that rotation cycle of the striped pattern coincides with exactly integer \( n \) fields \((n/60 \text{ s} \) for common NTSC video cameras) and so that rotational speed is kept constant. Though the brightness is used as representative in these discussions, the component of red, blue, and green, that is three primary colors, can be treated as same as in color images.

To detect the phase \( \phi \), we can integrate \( Y_{xy}[i](i = 0 \cdots n - 1) \) after multiplied with reference, \( \sin(2\pi i/n) \) and \( \cos(2\pi i/n) \) as follows:

\[
\begin{align*}
\bar{Y}_{C,xy} &= (1/n) \sum_{i=0}^{n-1} \sin(2\pi i/n) Y_{xy}[i] \\
&= (1/n) \sum_{i=0}^{n-1} \{aY_{0,xy} + Y_{B,xy}\} + (bY_{0,xy}/n) \sum_{i=0}^{n-1} \{\sin(2\pi i/n) \sin(2\pi i/n + \phi)\} \\
&= \{(aY_{0,xy} + Y_{B,xy})/n\} \sum_{i=0}^{n-1} \sin(2\pi i/n) \\
&+ (bY_{0,xy}/2n) \sum_{i=0}^{n-1} \{\cos(\phi) - \cos(4\pi i/n + \phi)\} \\
&= 0 + (bY_{0,xy}/2) \cos \phi = (bY_{0,xy}/2) \cos \phi. \tag{4}
\end{align*}
\]

As the same manner,

\[
\begin{align*}
\bar{Y}_{S,xy} &= (1/n) \sum_{i=0}^{n-1} \{\cos(2\pi i/n) Y_{xy}[i]\} \\
&= (bY_{0,xy}/2) \sin \phi. \tag{5}
\end{align*}
\]

We can obtain \( \phi \) from \( \bar{Y}_{C,xy} \) and \( \bar{Y}_{S,xy} \) at all pixel of the image.

In addition, we can also obtain ordinary image by summing up \( Y_{xy}[i] \).

\[
\begin{align*}
\bar{Y}_{xy} &= (1/n) \sum_{i=0}^{n-1} Y_{xy}[i] \\
&= (1/n) \sum_{i=0}^{n-1} \{aY_{0,xy} + Y_{B,xy}\} \\
&+ (bY_{0,xy}/n) \sum_{i=0}^{n-1} \sin(2\pi i/n + \phi) \\
&= \{aY_{0,xy} + Y_{B,xy}\} \tag{6}
\end{align*}
\]

The result \( \bar{Y}_{xy} \) is the mean of brightness without sinusoidal component.

Thus, we can obtain phase of each pixel and its original image.

### C. Use of higher harmonics

We assumed that a rotation cycle of the striped pattern coincides with period of image processing fields \( n \) in above discussion. We can also use higher harmonics of above fundamental wave. In case that we use \( s \)th harmonics of fundamental cycle, observed sequence of brightness is given by

\[
Y_{xy}[i] = Y_{0,xy}(a + b \sin(2\pi si/n + \phi)) + Y_{b,xy}. \tag{7}
\]

As same as (4), (5), we integrate \( Y_{xy}[i] \) after multiplied with reference, \( \sin(2\pi si/n) \) and \( \cos(2\pi si/n) \) as follows:

\[
\begin{align*}
\bar{Y}_{C,xy} &= (1/n) \sum_{i=0}^{n-1} \sin(2\pi si/n) Y_{xy}[i] \\
&= \{(aY_{0,xy} + Y_{b,xy})/n\} \sum_{i=0}^{n-1} \sin(2\pi si/n) \cos(2\pi si/n + \phi) \\
&+ (bY_{0,xy}/2n) \sum_{i=0}^{n-1} \{\cos(\phi) - \cos(4\pi si/n + \phi)\} \\
&= (bY_{0,xy}/2) \cos \phi, \tag{8}
\end{align*}
\]

\[
\begin{align*}
\bar{Y}_{S,xy} &= (1/n) \sum_{i=0}^{n-1} \{\cos(2\pi si/n) Y_{xy}\} \\
&= (bY_{0,xy}/2) \sin \phi, \tag{9}
\end{align*}
\]

where \( s \) must not be multiple of \( n \).

Higher harmonics can be obtained in two cases. One is to rotate the cylinder \( s \) times faster than the original case. This case is described in detail later. Another is to create striped pattern \( s \) times denser than original. The distinguishable range of phase becomes \( s \) times in this case though phase-connecting problem rises in common with phase-shift method. No large difference is found between range \([0, 2\pi]\) of fundamental case and range \([0, 2\pi s]\) in dense case reduced to a scale of \(1/s\) if we can obtain precise stripe pattern, though improvement in S/N is expected by repeated process.

### D. Simultaneous detection with frequency multiplex

The detection method shown above has strong frequency selectivity. This characteristic enables simultaneous detection by several pairs of light sources and cameras.

When we have two sets of equiphase planes from two light sources whose frequencies differ from each other, the brightness is given by

\[
Y_{xy}[i] = Y_{0,xy}(a + b_1 \sin(2\pi s_1 i/n + \phi_1)) + b_2 \sin(2\pi s_2 i/n + \phi_2), \tag{10}
\]

where \( s_1, s_2 (s_1 \ne s_2) \) are magnification of frequencies of two sets, \( b_1, b_2 \) are their amplitude of brightness, \( \phi_1, \phi_2 \) are phases, respectively.
The detection of component of frequency 1 requires references of \( \sin(2\pi s_1 i/n) \) and \( \cos(2\pi s_1 i/n) \) as follows:

\[
Y_{C1,xy} = (1/n) \sum_{i=0}^{n-1} \{ \sin(2\pi s_1 i/n) Y_{xy}[i] \}
\]

\[
= (1/n) \sum_{i=0}^{n-1} \{ \sin(2\pi s_1 i/n) (\alpha Y_{0,xy}) \}
\]

\[
= (1/n) \sum_{i=0}^{n-1} \{ \sin(2\pi s_1 i/n) \sin(2\pi s_1 i/n + \phi_1) \}
\]

\[
+ (b_1 Y_{0,xy}/n) \sum_{i=0}^{n-1} \{ \sin(2\pi s_1 i/n) \sin(2\pi s_1 i/n + \phi_1) \}
\]

\[
+ (b_2 Y_{0,xy}/n) \sum_{i=0}^{n-1} \{ \sin(2\pi s_1 i/n) \sin(2\pi s_1 i/n + \phi_2) \}
\]

\[
= (b_1 Y_{0,xy}/2n) \sum_{i=0}^{n-1} [\cos(\phi_1) - \cos(4\pi s_1 i/n + \phi_1)]
\]

\[
+ (b_2 Y_{0,xy}/2n) \sum_{i=0}^{n-1} \{ \cos(\phi_2 + 2\pi s_2 i/n - 2\pi s_1 i/n) \}
\]

\[
= (b_1 Y_{0,xy}/2) \cos \phi_1
\]

\[
+ (b_2 Y_{0,xy}/2n) \sum_{i=0}^{n-1} [\cos(\phi_2 + 2\pi(s_2 - s_1)i/n) \}
\]

\[
= (b_1 Y_{0,xy}/2) \cos \phi_1.
\]

\[
\sum_{i=0}^{n-1} \{ \sin(2\pi i/n + \phi) \}
\]

\[
= \sin(2\pi i + 2\pi i/n + \phi) = \sin(2\pi i/n + \phi).
\]

By observing through video camera, the images are same in case of \( s = 1 \) and \( s = n + 1 \), whereas human feel unwell in former case and perceive as simple light in latter case. This means that we can design system that works at low frequency for development and then speed up the rotation of striped cylinder \( n + 1 \) time for practical system. In addition, the light of \( s = n + 1 \) have actual frequency component higher than sampling frequency, that is video rate of 60 [Hz]. Of course, we need the camera with higher shutter speed that can follow the actual frequency of the light enough.

### III. TRIAL SYSTEM AND EXPERIMENTS

#### A. Experimental system

A trial system was developed to confirm above principle and effectiveness of the method. Figure 5 shows the experimental system. Outline of the system was shown in Fig. 5(a). It consisted of light source as shown in Fig. 5(b) (placed at top of the picture), a NTSC video camera (400,000 pixels, color CCD), motor controller (placed at side of light source), and processing computer (not in the

---

**E. Using sampling theorem to reduce flicker**

Conventional vision range sensors that use special light sources seem to make human feel unwell. Their lights always switch several light patterns or put strange patterns on measuring target. The sensors on robots in human living environment should be gentle to human in addition to safe to human. The method proposed in this paper can be gentle to human.

According to the discussion in section II-C, we can use harmonics of fundamental wave. If frequency of change of brightness is high enough than critical flicker frequency, the light from our source can be observed as non-flicker illumination.

In special case that frequency magnification ratio \( s \) is near to multiple of \( n \), for example \( s = n + 1 \), we can use effect of sampling theorem as follows:

\[
\sin(2\pi s_1 i/n + \phi) = \sin(2\pi(n + 1)i/n + \phi)
\]

\[
= \sin(2\pi i + 2\pi i/n + \phi) = \sin(2\pi i/n + \phi).
\]

By observing through video camera, the images are same in case of \( s = 1 \) and \( s = n + 1 \), whereas human feel unwell in former case and perceive as simple light in latter case. This means that we can design system that works at low frequency for development and then speed up the rotation of striped cylinder \( n + 1 \) time for practical system. In addition, the light of \( s = n + 1 \) have actual frequency component higher than sampling frequency, that is video rate of 60 [Hz]. Of course, we need the camera with higher shutter speed that can follow the actual frequency of the light enough.
The lighting unit in Fig. 5(b) consisted of a stepper motor (400 pulse/rev), a striped cylinder made of stripe-printed OHP transparency sheet and acrylic cup, and a halogen lamp for headlight of a car. The motor rotated at speed of 60 steps per second synchronized to V-sync of video signal from camera.

The function of transparency was given by

$$\alpha(\psi) = 0.5 + 0.25 \cos(5\psi) + 0.25 \cos(50\psi),$$

which means combination of wave of 5 cycles and 50 cycles as shown in Fig. 5(c). One-fifth of the pattern, that is combination of 1 cycle and 10 cycles, was used as a set of images for the process. Therefore, 80 fields ($= 400/5$) of images were taken in 1.33[s] ($= 80/60$), and measuring area were limited to 72 [deg] fan-shaped area from the light source.

The video images were captured to a PC of Pentium4 2.8 [GHz]. The operating system of the PC was Linux 2.4 (distribution of KNOPPIX EduTG 0.9). All the processes were implemented with software written in C++.

The summing-up calculation (8) and (9) of two references (1 cycle and 10 cycles per the period) were done in each interval between capturing frames (two fields were obtained as a frame simultaneously via Video for Linux), which took about 5 [ms]. The finalized part that calculate phase of each pixel, convert to distance, and display results were processed after 80 images had been obtained, which took about 50 [ms]. A simple interpolation converted phase to distance after a set of two phases on each pixel was measured on two reference planes (that decided near and far limits).

### B. Experimental results

Figure 6 shows three of experimental results obtained by above system. Each result contained four images of 320 × 240 resolution. The top-left image showed distance of each pixel (range image). Darker pixel indicates nearer point was found at that pixel and brighter pixel means farther point (black and white pixels are near and far reference point, respectively). The bottom-left image is contour of distance that is derived from range image. The top-right image is normal image obtained with equation (6) simultaneously with range image. The bottom-right image is made from normal image and range image. The image was rotated a little around actual vertical axis by using distance of each pixel. The black region indicates the unknown pixels that are mainly back part of objects.

These results showed measurement ability of this method though the used system was incomplete one. We could obtain detailed range image of complex objects including repeated pattern of the basket. The outline of each object was easily recognized in range object and contour image. We can also find the shade of object itself under handle of the cup and around the basket, which were indicated black region in range image. These shades will disappear if we use more than one light source.

In addition, the technique reducing flicker described in section II-E was also tested with a new developing system that used DC servomotor that rotated much faster. The camera detected slow movement of stripes on the...
measuring targets whereas we could perceive as normal non-flicker light.

IV. CONCLUSION

A new approach of 3D vision range sensor was proposed in this paper. The method used a special light source that projects planes of lights whose intensity changes as sinusoidal function of time and have individual phase. Distance to each point on video image could be obtained from the phase that was observed at each pixel.

The basic principles and expansion for practical system were described, which needs lower calculation power than traditional stereo matching calculation. The method also considered the feeling of human who stays with the system so that the system can be applied to mobile robots in human living environment.

The experimental results were brought by trial system. They showed that effectiveness of proposed method. This method could obtain distance to each point in camera image.

A new practical system is under development. It uses a DC servomotor with optical encoders and a cylinder designed and manufactured precisely. The hardware processing system is also planed to realize the system that is more compact with less power consumption for mobile robots.

ACKNOWLEDGMENT

The author would like to thank engineers of campus workshop of Tohoku Gakuin University in creating practical experimental system.

REFERENCES