

Modified time of flight camera 3D-images improving method

D. FALIE, L. C. CIOBOTARU^{a*}

Politehnica University of Bucharest, Romania

^aNational Institute of Lasers, Plasma and Radiation Physics, Magurele, Romania

The 3D shape image of high contrast objects obtained with the Time of Flight cameras presents significant distortions. These distortions are caused either by the scattered light inside the camera body or by the indirect (diffuse) light illuminating the objects in the scene. The modified Time of Flight (ToF) camera corrects the 3D image by illuminating the scene with two consecutive light intensities. In one frame the scene is normally illuminated and in the consecutive frame an additional spotlight is added. A correction vector is computed using these two images. The improved 3D image is computed by subtracting this vector from the original vector image which leads to a significant improvement of 3D images. Thus, the modified ToF camera becomes easily to calibrate and its use could be extended in many industrial applications.

(Received July 28, 2009; accepted February 02, 2010)

Keywords: 3D image, Time of flight camera, Image corrections

1. Introduction

The 3D Time of Flight (ToF) movie camera are produced by MESA Imaging, PMD, Canesta and 3DV Systems. All these cameras are small and can be connected to a PC by an USB interface. The next cameras generation is expected to be mainly 3D cameras and many possible medical applications could be developed in the future for these devices [1].

The ToF camera gives two images; one is the *distance image* and the other is the usual black and white *amplitude image* or a *color image* in the 3DV Systems' camera. In this experiment it was used the model SR3000, manufactured by MESA, with the following characteristics: distance and amplitude images size are 176 by 144 pixels and the typical frame rate is of 25 frames per second. This camera illuminates the scene with an infrared light source which is amplitude modulated with a high frequency signal. The wave length of the modulation frequency for the SR3000 camera is 15 m corresponding to a frequency of 20 MHz.

Each camera's pixel i measures the phase difference φ between the modulation signal and envelope of the received light signal. The distance $d(n_i, n_c)$ from an object to its image, namely the pixel with the row number n_i and column number n_c , is computed with relation (1), where $\varphi(n_i, n_c)$ is the measured phase difference, c_0 is the speed of light and f the modulation frequency [2,3].

$$d(n_i, n_c) = \frac{c_0 \cdot \varphi(n_i, n_c)}{4 \cdot \pi \cdot f} \quad (1)$$

The ToF camera working principle is presented in Fig. 1.

The real distance to an object is, in fact, the distance measured by the camera plus a multiple of $\lambda/2$. In this case, the maximum non-ambiguity distance range is equal to $\lambda/2$ which means that for the modulation frequency of 20 MHz the calculated value is $\lambda/2=7.5$ m. It is preferably than no object in the scene will be placed further than a distance equal with $\lambda/2$.

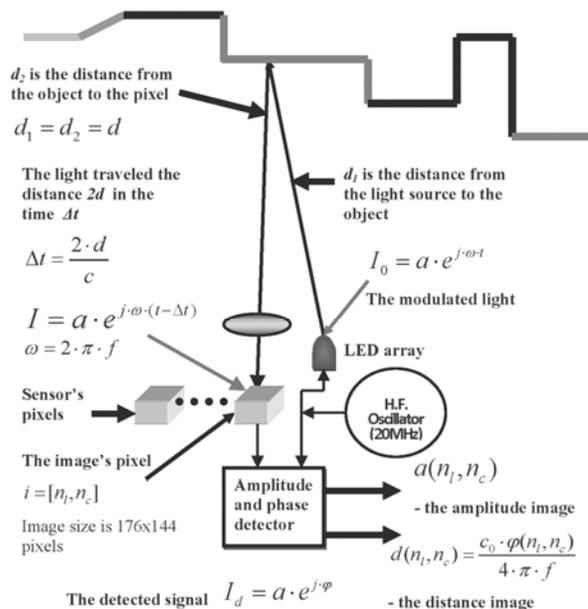


Fig. 1. The working principle of the ToF camera.

The great part of the ambient light is filtered out by a narrow band infrared filter. This filter permits to pass through only the light with the same wavelength as the one

emitted by the camera, namely with $\lambda = 850$ nm. The ambient light is even more attenuated because camera's pixels detect the amplitude modulated light with a specific frequency. The amplitude image represents only the image produced by the reflected light emitted by the camera's illuminating source. This reflected light decreases with the distance square to the illuminated objects and for this reason the objects with similar reflectivity placed at different distances appear to be also different.

The main advantage of the ToF camera is the fact that the *distance image* and the measured distance by each pixel is given by the phase difference between the reflected light signal and the modulation signal. The principal problem is that not all the light reflected by an object is produced by the direct light emitted camera ; a small fraction of it is an indirect light. This indirect reflected light, which comes from other objects prior illuminated by the camera's lighting source, has traveled across a longer path than the direct reflected light. In these circumstances the measured distance will be affected of errors due this diffuse light, which illuminates all the objects placed in the scene.

The diffuse light which suffered multiple reflections always gives a greater distance than the real one. This means that the measured distance by the camera will be greater because a small fraction of the received light has already suffered multiple reflections [4-6]. Generally, the introduced error is small and unimportant for example in gaming but in industrial applications the distorted 3D image implies unacceptable quantitative errors. This phenomenon affects almost equally the black and/or white objects and the 3D image of the strong contrast object is not significantly distorted. The measured distance to the objects is longer than the real one and this error is the same for the non-glossy (mate) objects white or black. The 3D image of objects with glossy surfaces has important distortions caused by the diffuse light. The extreme case is the measured distance to a mirror when this is the distance to the object whose image is reflected by the mirror and not the distance where the mirror is placed. A real surface is more or less glossy, and the measured distance to it with a ToF camera also depends on its orientation. The measured distance to a glossy object is higher than that to a mate one, and generally its 3D image is distorted, for example, the 3D image of a cube with glossy surfaces.

Inside the camera body not all the incoming light is absorbed by the sensor and a part of it is reflected. After multiple reflections, a fraction of this light is detected by other pixels, and it produces distance errors. In this case, the measured distance can be smaller or greater than the actual one depending on the traveled distance by the perturbing light. The scattered light inside the camera body usually affects the image of the black objects but does not emphasize the objects' glossy surfaces.

The errors produced by the diffuse light in the scene

or the scattered light inside the camera body obviously cannot be completely eliminated, and any correction method is benefic for the further development of this new type of camera. In many applications, the main concern is not the distance errors, especially in the movie camera case, but the distance image distortions. For instance, the measured distance error to a person is only of a magnitude of a few cm, but the black strips of the T-shirt appears to be 10 cm further than the white ones, and such distortions are unacceptable (Fig. 3).

The main purpose of this paper consists in proposing a method in order to diminish these distance image distortions.

2. Theoretical principles of the 3D image improving method

As it was already mentioned, the distance image is affected by a perturbing light signal produced by the light scattered both in the scene and inside the camera body. These phenomena could create problems in camera calibration because the distance measured depends equally on the room and on the objects in the room [6]-[11]. In order to minimize light reflections, the walls of the calibration room are painted in black. The scattered light signal generally has a slow spatial variation, but it thins are different when the objects are moving in the scene. In the 3D movie, even fixed objects seem to move when other objects move. To correct the 3D image, we used the method proposed in [8]. We added an LED similar to those used in the camera lighting source. The illumination angle of this LED is reduced to about 10 by 10 pixels by a lens. This additional LED is powered similarly as the camera's LED, but while in one frame it is on, in the consecutive frame it is off.

The ToF camera illuminates the scene with 55 LED; the command signal for the additional LED is extracted with a photo diode placed in front of one the cameras LED as in the Fig. 2. The amplified signal from the photo diode is used to power the additional LED through a logic circuit "and"- type. The camera's LED are powered from the high frequency (H.F.) oscillator only during the exposure time, after this until the next frame, the LED are switched off. After the exposure time, the camera processes and transfers to the PC the acquired data and during this time the camera's LED are switched off.

By integrating the amplified signal from the photo diode is obtained the exposure time pulses which are then divided by 2. Because the divided signals are two time longer it is necessary to short the duration of the exposure time with the logic circuit "and". The output of this gate is on only when the exposure time is on, this signal pass through another gate type & and the output of this gate is

on only when the camera's LED are on. These final signal powers the additional LED only half of frames.

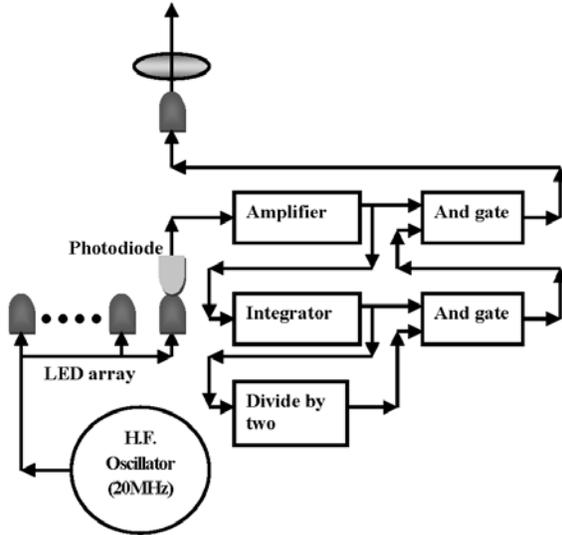


Fig. 2. The connection diagram of the additional LED.

Here is a brief description of the measured distance correction method: the detected signal is divided into two components; one component $I_d(i)$ is the signal produced by the light emitted by the camera, reflected once and then detected. The other component $I_p(i)$ is the sum of all other signals, and it is mainly produced by the light emitted by the camera and reflected twice or more times before it is detected. For each pixel i , the detected signal I_m is the sum of these two components:

$$I_m(i) = I_d(i) + I_p(i) = a_d(i) \cdot e^{j\varphi_d(i)} + a_p(i) \cdot e^{j\varphi_p(i)} \quad (3)$$

For the regions that are illuminated by the spotlight, there are two similar equations for two consecutive frames. In the frame where the spotlight is added, the pixels receive k times more light than in the previous image where the spotlight is off. It can be approximated that, in these two consecutive images, the perturbing component $I_p(i)$ is the same. Such an approximation will produce a small error if the difference between the perturbing components is about 1%. This happens if the spotlight illuminates only 1% of the pixels with the same power as the original light. In this case these pixels will receive two times more light, but this light will contribute only 1% to the perturbing signals of the scene.

For each pixel i in these regions, we have two equations:

$$I_{m1}(i) = I_{d1}(i) + I_p(i) = a_{d1}(i) \cdot e^{j\varphi_d(i)} + a_p(i) \cdot e^{j\varphi_p(i)} \quad (4)$$

$$I_{m2}(i) = I_{d2}(i) + I_p(i) = a_{d2}(i) \cdot e^{j\varphi_d(i)} + a_p(i) \cdot e^{j\varphi_p(i)}, a_{d2}(i) = k \cdot a_{d1}(i) \quad (5)$$

From the relations (4) and (5) it can be computed the direct $I_d(i)$ and the perturbing $I_p(i)$ components of the signal:

$$I_d(i) = \frac{1}{k-1} (I_{m2}(i) - I_{m1}(i)) \quad (6)$$

$$I_p(i) = \frac{1}{k-1} (k \cdot I_{m1}(i) - I_{m2}(i)) \quad (7)$$

For a given value $I_p(i)$ of a pixel, it can be computed $I_d(i)$ in the neighboring regions by subtracting the determined perturbing component $I_p(i)$ from the measured values. The $I_p(i)$ component has a slow spatial variation which it is not constant, and the correction is acting effectively only in a very limited neighborhood.



Fig. 3. The amplitude image obtained by the ToF camera.

In the scene image, the spotlight is the little white square in the upper part of the picture in Fig. 3, which is the LED image focused by a lens. The calibrated $k(i)$, plotted in Fig. 4, shows significant variations; for this reason, when $I_p(i)$ is computed, it is used only the eight pixels with the highest values of $k(i)$. The final value of I_p is the vector average over the eight values $I_p(i)$ computed with the eight calibrated constants $k(i)$.

The calibrated $k(i)$ is, in fact, the ratio of the two amplitude images with and without the additional spot light in the absence of any scattered light. Such a situation where is no scattered light exist in special laboratory setup, for example, when all the objects are absolutely black except a small white screen illuminated by the additional LED. In this situation, the white screen is illuminated only by the direct light because all other objects are black, and they do not reflect any other light. Such a situation is approximately accomplished in the laboratory when the

ToF cameras are calibrated. Better conditions exist if the camera is calibrated in a free space during the night. The reflected light decreases with the square root of the distance to the reflecting object, if it will place the white screen at 1 m distance and all other objects are placed farther than 33m- then the scattered light can be neglected.

From the relation (6) results the fact that the corrected distance is the phase of $I_d(i)$ and it is independent on $k(i)$; for those zones illuminated by both light sources, it can be computed the corrected distance also from relation (6) and is not necessary to compute $I_p(i)$. In the neighborhood region, the corrected distance is computed with the relation:

$$I_d(n_l, n_c) = I_m(n_l, n_c) - I_p \quad (8)$$

It is assumed the fact that $I_p(n_l, n_c) \cong I_p$ and from the experimental results it can be observed that this generally happens in a large neighborhood.

The *distance image* or the 3D image shows distortions of the objects with high contrast, and in order to put in evidence such distortions it were placed in the scene black square tags on white surfaces. The spotlight's optical axis is slightly different from the camera's optical axis, and the projection angle is also not identical with the viewing angle of the camera's lens. These facts imply that, for various distances from the camera to the object illuminated by the spotlight, the calibrated pixel i in the calibration setup does not correspond to the same pixel in the scene image. To identify the corresponding calibrated pixels, it has been established a spotlight zone in which it can be computed the ratio of two consecutive frames. The eight pixels with the highest values of the image ratios correspond to those eight pixels with the highest $k(i)$ values.

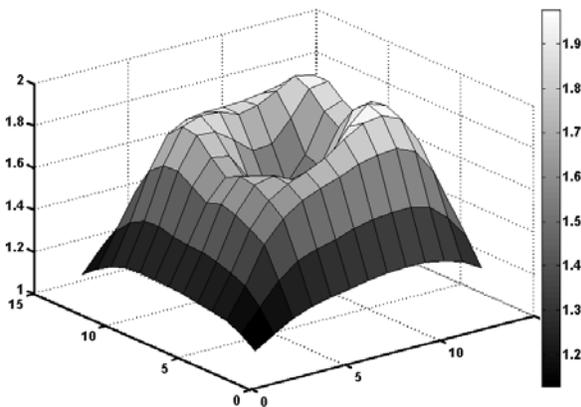


Fig. 4. The 3D representation of the k coefficient (the gray scale and the vertical axes represent the k values and the horizontal axes represent the lines and columns of the spotlight image).

3. Results and discussion

By applying the above method to the scene from the Fig. 3, it can be obtained significant 3D image improvements.

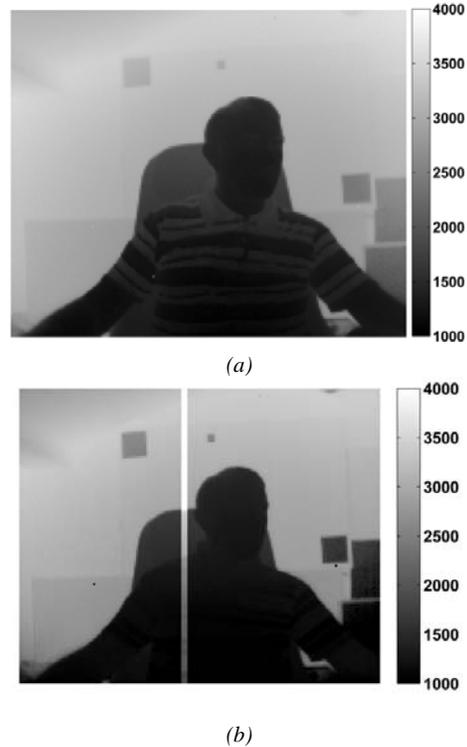


Fig. 5. (a) The uncorrected distance image; the measured distance to the black and white part of an object is different; (b) The improved distance image (the gray scale represents the measured distance in mm).

The original distance image is shown in the Fig. 5 (a), where it can be observed that the measured distance to the T-shirt strongly depends on the brightness of the object and for reason is highly distorted. In the improved image, shown in the Fig. 5 (b), the distance image in the region of the T-shirt is significantly improved. The distance image improvement decreases for regions situated further from the pixels where $I_p(i)$ was calculated (the white square shape region above the head of the person from the Fig. 3). The image improvement decreases with the distance from the white spot which can be observed by comparing the distance values for the column highlighted in Fig. 5 (b) (along the vertical white line). In Fig. 6, the distance in mm is plotted on the vertical axis, and the pixel number along the vertical white line of the Fig. 5 (b) is plotted on the horizontal axis. The white spotlight is situated around the pixel 30. It is remarkable a steady improvement along extended regions, and only at the far bottom of the image it can be noticed an overcorrection. On the corrected distance image, around pixels 144 to 110, the black stripes of the T-shirt appear closer to the camera while in the original image they are farther away.

Between the uncorrected and corrected curves in Fig. 6 is a constant gap because the diffuse light which illuminates the objects in the scene has passed a longer

distance than the direct reflected light. In this case, all the objects in the scene will appear at a longer distance than they are in reality and for this reason, there is a gap between the plot of the corrected and uncorrected distances in Fig. 6. For distances longer than 2500mm the gap is smaller because the phase of the perturbing signal is closer to that of the direct signal. For closer distances - around 1500mm - the error gap increases because the phase difference between the direct and perturbing signals becomes greater. The error produced by the perturbing signal seems to decrease for closer distances around 1000 mm caused by the amplitude increase of the direct signal; the amplitude of the reflected light decreases with the square root of the distance to the object.

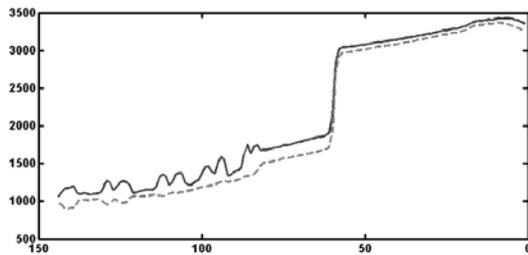


Fig. 6. The measured distance (solid line), and the corrected distance of the improved image (dotted gray line); the vertical axis is the distance in mm and the horizontal axis is the pixel number.

Comparing a 3D image of the T-shirt (Fig. 7 (a)) with the improved one (Fig. 7 (b)), besides the significant improvement it is obvious that the bottom part of the T-shirt is distorted in an opposite direction on the improved image. In Fig. 7 (a) and (b) the z scale of the images was limited to the interval [1000mm, 2000mm] in order to put in evidence in the region of the T-shirt the 3D image distortions.

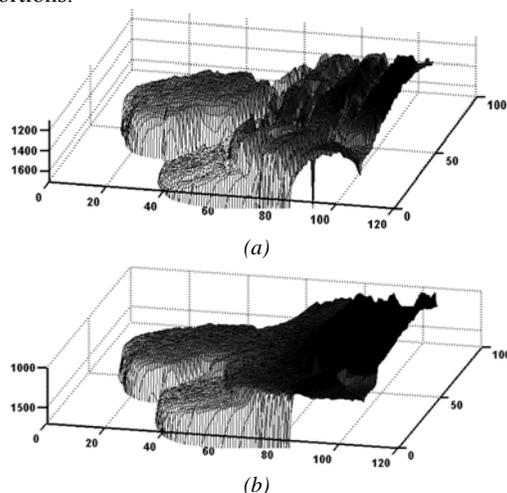


Fig. 7. (a) The 3D image of the T-shirt (the vertical axis represents the distance in mm, and the horizontal axes show the line and column numbers on the image); (b) The improved 3D image of the T-shirt (the vertical axis represents the distance in mm and the horizontal axes show the line and column numbers on the image).

4. Conclusions

The modified Time of Flight camera gives 3D improved images using the proposed method. These results could be very useful in industrial applications where higher distance accuracy is required. The camera can be easily calibrated in any laboratory.

To improve the entire image, more spotlights must be added. Instead of LED is better to use laser diodes, and the four thin horizontal line shaped lasers (two lines in a frame and other two interleaved lines in the consecutive frame) will be sufficient to improve the whole image.

It is important to notice that the perturbing signal is computed for each frame using the two consecutive images but the frame rate is not affected.

Acknowledgements

This research was accomplished in the framework of the European Project "Action Recognition and Tracking based on Time-of-Flight Sensors" (ARTTS). The ARTTS project is funded by the European Commission (contract no. IST-34107/ Information Society Technologies - the 6-th Framework Program).

The present paper reflects only the views of the authors and the Commission cannot be held responsible for any use which may be made of the information contained herein.

References

- [1] D. Falie, L. David, M. Ichim, in: Int. Symp. on Sig., Circ., and Systems (ISSCS), Iassy, 1, IEEE Cat. Nr. CFP09816-PRT, **1**, 45 (2009).
- [2] Z. Xu, R. Schwarte, H. Heinol, B. Buxbaum, T. Ringbeck, in: Proc. Int. Conf. on Mechatron. & Machine Vision, 259, 1998.
- [3] T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, N. Blanc, Proc. SPIE, **5249**, 634 (2003).
- [4] S. Oprisescu, D. Falie, Mihai Ciuc, Vasile Buzuloiu, ISSCS, Iassy, 2212007.
- [5] R. Lange, Ph.D. thesis, University of Siegen, 2000.
- [6] MESA Imaging SA, SwissRanger SR-3000 Manual Version 1.02 October, 2006.
- [7] D. Falie, V. Buzuloiu, 2008 Microwaves, Radar and Remote Sensing Symposium (MRRS-2008), September 22-24, Kiev, Ukraine, 79, 2008.
- [8] D. Falie, V. Buzuloiu, IEEE Workshop on Imaging Systems and Techniques (IST2008), September 10-12, Chania, Crete, Greece, 123, 2008.
- [9] D. Falie, V. Buzuloiu, Fourth European Conference on Circuits and Systems for Communications (ECCSC'08), July 10-11, Bucharest, Romania, 121, 2008.
- [10] D. Falie, V. Buzuloiu, Fourth European Conference on Circuits and Systems for Communications (ECCSC'08), July 10-11, Bucharest, Romania, 125, 2008.
- [11] D. Falie, Proceedings of the SPIE, **7156**, 715615-715615-11.

*Corresponding author: cataliniobotaru@yahoo.com