3D Image Correction for Time of Flight (ToF) Cameras

D. Falie
The Image Processing and Analysis Laboratory, Politehnica University of Bucharest, Romania.

ABSTRACT

The distance image from the ToF camera is affected by errors mainly due to multiple light reflections inside the camera body and by the indirect light illuminating the objects in the scene. These errors can't be avoided and in particular cases can be as great as 50%. The method presented corrects the measured distance errors to objects labeled with a known strong contrast tag. A proposed algorithm determines the integral perturbing light signal and by subtracting it from the measured one we compute the (actual) corrected 3D image. Because the perturbing signal has a slow variation along the image it is computed only in the region of the tag with known contrast but the correction is performed in a wider region. An extension of this method to the entire 3D image can be implemented using amplitude modulated structured light. The method can be useful to other distance measuring devices using the same (ToF) principle.

Keywords: 3D, camera, image, corrections

1. INTRODUCTION

TOF cameras are a new type of 3D cameras which have their own illumination source so they can work even when there is no other light source. Each pixel of the camera measures the incoming reflected light and the distance to the objects in the scene. This information is obtained by using amplitude modulated light. The amplitude and the phase of the modulating wave are detected by a phase detector. In this way only the modulated light is detected and the background light is rejected [1].

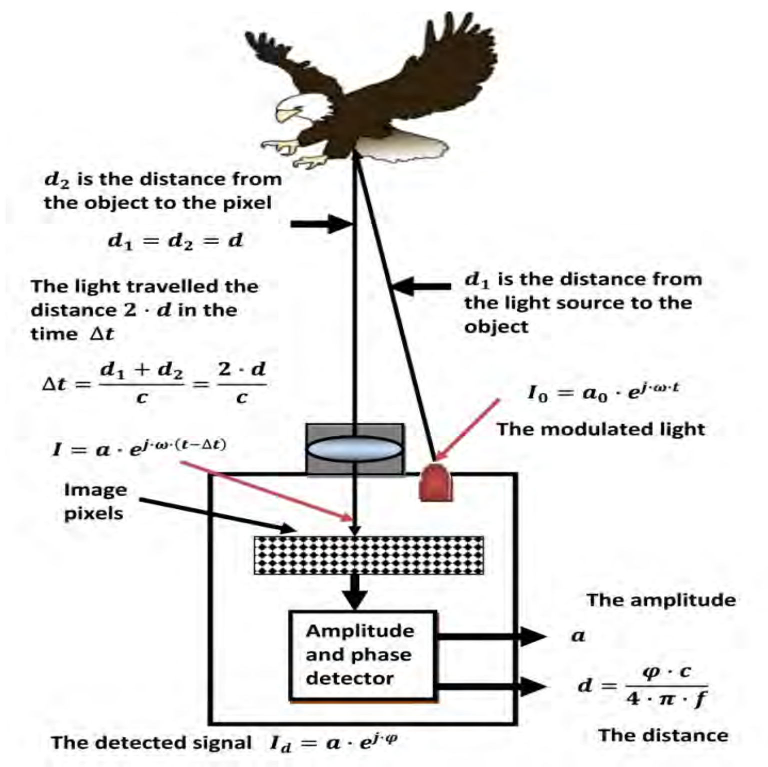

Fig. 1. The principle of the Time of Flight camera.

There is also a disadvantage of this illumination mode the intensity of the reflected light decrease with the square root of the distance to the objects and the amplitude image has a high dynamic range. A correction of the amplitude image is to multiply it with the square of the distance image [2].

The camera acquires two images at the same time: one is the classical amplitude (brightness) image and the second is the “distance image”.

The work principle of the ToF camera is represented in Fig. 1. The distance to the object is given by relation (1) where the phase difference between the emitted and reflected modulated light signal is φ, *c* is the speed of light and *f* the modulation frequency.

$$d = \frac{c \cdot \varphi}{4 \cdot \pi \cdot f} \tag{1}$$

The in-phase detection of the amplitude and phase of the modulation signal is achieved as follows: the camera’s pixels integrate the electrical charge produced by the incoming light during a *T/2* time interval (*T=1/f*) and in the next T*/4* time interval the accumulated charge is transferred to the first storage cell (sample no. 1): in the next 3T*/4* time interval the same process is repeated and the accumulated charge is transferred to another stoage cell (n0. 2). This process is repeated four times and each time the accumulated charge is transferred to a different storage cell, so the samples no. 1, 2, 3 and 4 are formed. After that the same process restarts and repeats the 4 sample cycle up to the end of the exposure time t_{ex}. The time needed to detect the incoming light and to transfer the (detected) electrical charge to the appropriate sample is *T/2+T/4= 3T/4* and the time cycle for the four samples is *3T*. During the exposure time t_{ex} all this entire process is repeated $t_{ex}/(3\cdot T)$ times.

Suppose now that the light signal emitted by the camera is sinusoidal of the form $I_e = a_e\left(m + \sin(\omega \cdot t)\right)$, where $m \geq 1$. The detected (measured) signal is also sinusoidal but with a different phase:

$$I_m = a_m\left(m + \sin(\omega \cdot t + \varphi)\right) \tag{2}$$

The four samples of the detected light are:

$$S_i = \frac{t_{ex}}{3 \cdot T} \cdot \int_{(i-1)\frac{3 \cdot T}{4}}^{i\frac{3 \cdot T}{4}} a_r\left(m + \sin(\omega \cdot t + \varphi)\right) \cdot dt\,, \quad \iota = 1,2,3,4 \tag{3}$$

which gives us

$$S_1 = \frac{a_r \cdot t_{ex}}{3}\left(\frac{m}{2} + \frac{1}{\pi}\cos(\varphi)\right), \quad S_3 = \frac{a_r \cdot t_{ex}}{3}\left(\frac{m}{2} - \frac{1}{\pi}\cos(\varphi)\right) \tag{4}$$

$$S_2 = \frac{a_r \cdot t_{ex}}{3}\left(\frac{m}{2} + \frac{1}{\pi}\sin(\varphi)\right), \quad S_4 = \frac{a_r \cdot t_{ex}}{3}\left(\frac{m}{2} - \frac{1}{\pi}\sin(\varphi)\right) \tag{5}$$

From these four samples we get the two important values of the detected light signal I_x and I_y:

$$I_x = S_1 - S_3 = \frac{2a_r \cdot t_{ex}}{3\pi}\cos(\varphi)\,, \quad I_y = S_2 - S_4 = \frac{2a_r \cdot t_{ex}}{3\pi}\sin(\varphi) \tag{6}$$

The amplitude a_m and the phase φ of the detected light are:

$$a_r = \frac{3\pi}{2t_{ex}}\sqrt{I_x^2 + I_y^2} = \frac{c_d}{t_{ex}}\sqrt{I_x^2 + I_y^2} \tag{7}$$

$$\varphi = \pi + \arctan\left(\frac{I_y}{I_x}\right) \tag{8}$$

where c_d is a detection constant. The electrical charge produced by the incoming light in a pixel during the integration time $T/2$=25 ns (in the case when the modulation frequency is 20 MHz) is at most 1(one) electron!
The maximum non-ambiguous distance range of the ToF camera is:

$$d_{\max} = \frac{c}{2f} = \frac{\lambda}{2} \tag{9}$$

The measured distance ambiguity can be resolved by measuring the distance with two different frequencies f_1 and f_2. In a frame we measure the distance with one frequency and in the next frame with the other frequency and so on. In the non-ambiguity range of the camera the measured distance is the same but outside this region the measured distances are different and this helps us to determine the correct distance.

We have the relations:

$$d_1 = d_{m1} + n_1 \cdot \frac{c}{f_1} = d_{m1} + n_1 \cdot \lambda_1, \quad d_2 = d_{m2} + n_2 \cdot \frac{c}{f_2} = d_{m2} + n_2 \cdot \lambda_2 \tag{10}$$

where d_1, d_2 (d_1=d_2) represent the distances to the object, d_{m1}, d_{m2} are the measured distances with the corresponding frequencies f_1, f_2 and λ_1, λ_2 are the corresponding wavelengths, and n_1, n_2 are two integers which we must find in order to resolve the ambiguity. If we subtract the above relations we have:

$$d_{m1} - d_{m2} = n_2 \cdot \lambda_2 - n_1 \cdot \lambda_1$$

For the first ambiguity intervals $n_1 = n_2$ or $|n_1 - n_2| = 1$. If $\lambda_1 < \lambda_2$ the values of n_1, n_2 and the actual distance d are:

$$n_1 = n_2 = \mathrm{int}\left(\frac{d_{m1} - d_{m2}}{\lambda_2 - \lambda_1}\right); \qquad d = \frac{d_{m1}\cdot\lambda_2 - d_{m2}\cdot\lambda_1}{\lambda_2 - \lambda_1}; \text{ for } d_{m1} - d_{m1} \geq 0 \tag{11}$$

$$n_1 = \mathrm{int}\left(\frac{d_{m1} - d_{m2} - \lambda_1}{\lambda_2 - \lambda_1}\right); \qquad d = \frac{d_{m1}\cdot\lambda_2 - d_{m2}\cdot\lambda_1 + \lambda_1\cdot\lambda_2}{\lambda_2 - \lambda_1}; \text{ for } d_{m1} - d_{m1} < 0 \text{ and } n_2 = n_1 - 1 \tag{12}$$

In these relations "*int*" means round to the nearest integer.

1.1 Distance image errors produced by scattered light.

The distance measured by the ToF camera is affected by an unavoidable error. All the objects in the scene are illuminated not only by the direct light emitted by the camera but also by diffuse (scattered) light produced. The camera's modulated light is specular or diffuse reflected by the objects in the scene and after such multiple reflections all the objects are also illuminated by this indirect (diffuse) light. The detected light is in fact the sum of the direct reflected light and all indirect lights signals reflected. The phase of the indirect light signal is different from that of the direct one and the measured distance will be affected by this error. Unfortunately a similar phenomenon is produced inside the camera body. The incoming light is not totally absorbed by the chip and a fraction of this light is reflected. After multiple reflections by the lenses, I.R. filter and the walls of the cameras, a fraction of this light is added to the direct light and detected.

This phenomenon is well known and the users are advised to avoid some situations when large errors can be produced. Two of such situations are represented in Fig. 2 (reproduced from [3]). In the first situation the measured distance to the wall "A" is greater than the real one because the indirect light reflected by wall "B" is added to the direct light reflected by wall "A". In the second situation the measured distance to the dark object "Target 2" is smaller than the real one due to the light coming from the closer bright object "Target 1". The lenses are anti-reflex coated and the inner walls of the camera are painted in black so the indirect light signal is much less than 1% of the direct one. However, there are situations when the direct light reflected by a dark object is 100 times smaller than that from a closer bright object.

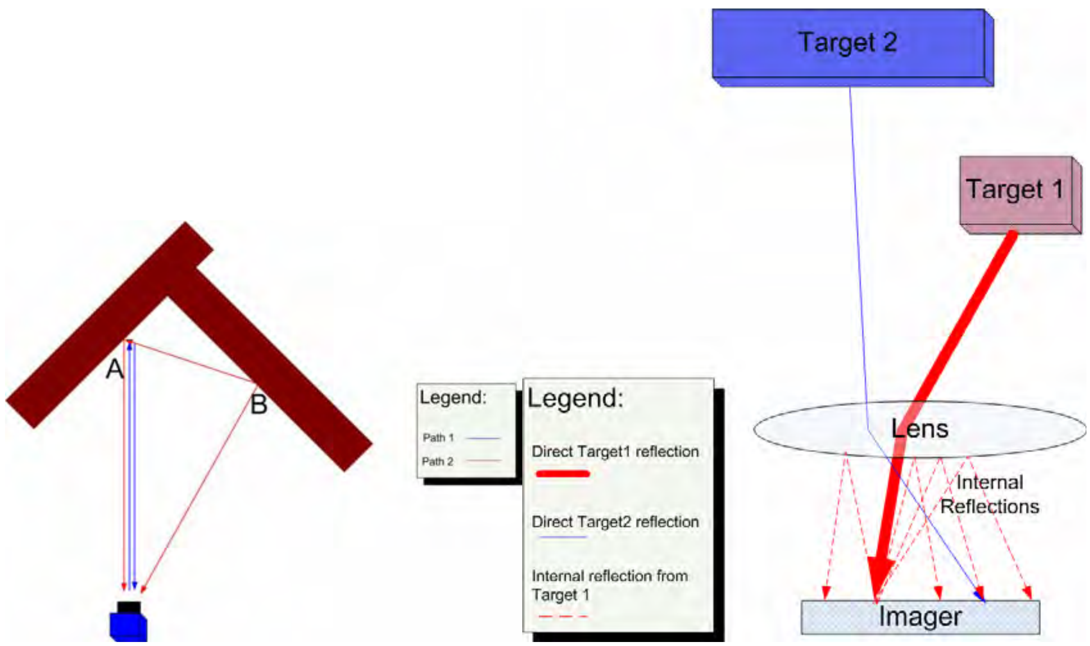

Fig. 2 Measured distance errors produced by diffuse light [3].

1.2 Correction of the errors produced by scattered light.

The light signal I_m detected by a pixel is the sum of the signals produced by the direct light reflected, I_d and the signal produced by the integral (sum) of all others unwanted (scattered) light signals, I_p. The background light or the natural light reflected by objects does not affect I_m. This is affected only by the amplitude modulated light emitted by the camera. For each pixel we have the relation:

$$I_{m1}(i) = I_d(i) + I_{p1}(i) = a_d(i) \cdot e^{j\varphi_d(i)} + a_{p1}(i) \cdot e^{j\varphi_{p1}(i)} \quad (13)$$

The distance measured by pixel i can be corrected if a structured light is used. In one frame the scene is illuminated as usual and in the consecutive frame a structured light is added. The signal detected by pixel i with the structured light added is:

$$I_{m2}(i) = k \cdot I_d(i) + I_{p2}(i) = k \cdot a_d(i) \cdot e^{j\varphi_d(i)} + a_{p2}(i) \cdot e^{j\varphi_{p2}(i)} \quad (14)$$

When the added structured light illuminate the region of pixel i it will increase the amplitude of the direct light signal reflected by a known factor k. If the power of the structured light added is much less (1/10 up to 1/100) than that of the original light we can consider that the perturbing signals are the same in both situations $I_{p1} = I_{p2} = I_p$.

Relation (13) and (14) form a system of two equations with two unknowns I_p and I_d.

$$I_d(i) = \frac{1}{k-1}\left(I_{m2}(i) - I_{m1}(i)\right) \tag{15}$$

$$I_p(i) = \frac{1}{k-1}\left(k \cdot I_{m1}(i) - I_{m2}(i)\right) \tag{16}$$

From (15) we observe that it is not necessary to know factor k in order to compute the corrected distance (φ_d). Because the scattered light has a smooth spatial variation knowing the value of I_p at a point will help correct the distance to the neighborhood points:

$$I_d(i) = I_m(i) - I_p \tag{17}$$

The proposed correction can be performed in real time because the detection algorithm (6) gives us the vector components I_x and I_y of the measured signal $I_m(i)$. With these components we compute (15), (16), (17) and then the amplitude and distance images. The amplitude image is also corrected with (15).

The need to quantitatively measure the light reflected by an object may appear in many industrial and scientific applications and with this method we can eliminate the influence of the scattered light.

1.3 The correction of errors produced by the light scattered inside the camera body.

This method corrects only the errors produced by the light scattered inside the camera body but it is useful in many applications to measure light scattered inside the camera body.

The image correction method is similar with the one presented above. We split the perturbing signal I_p in two components: one is the perturbing signal I_{pi} produced by the scattered light inside the camera body and the other is the perturbing signal I_{po} produced by the scattered light outside the camera body. The signal detected by the pixel i_1 is:

$$I_m(i_1) = I_d(i_1) + I_{po}(i_1) + I_{pi}(i_1) \tag{18}$$

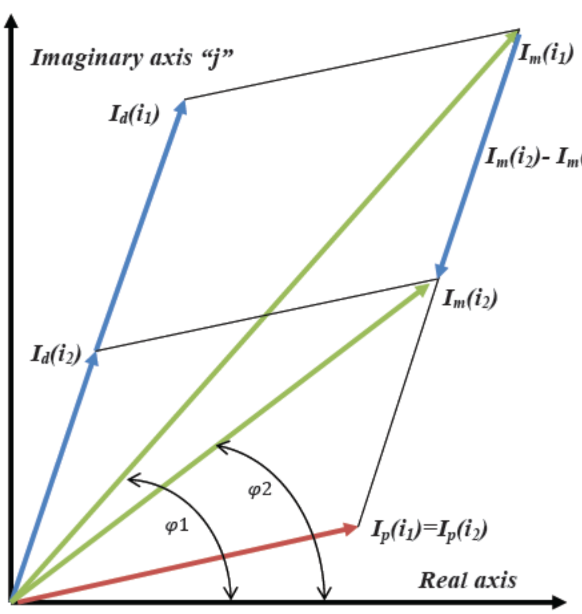

Fig. 3. In the presence of scattered light the measured distance to two neighborhood regions of an object - one white and the other black - are different but the vector difference of these two measurements has the correct distance.

We put a high contrast tag in the scene (for example a small sheet of paper half white and half black). We now suppose that we know the contrast factor k_t between the white and black parts of the paper (k_t is the ratio of the light reflected by the white and black part of the tag). We choose two neighborhood pixels one in the black part of the tag i_1 and the other in the white part i_2. The signal measured by these two pixels is given by (14) and (15).

$$I_m(i_2) = I_d(i_2) + I_{po}(i_2) + I_{pi}(i_2) = k_t\left(I_d(i_1) + I_{po}(i_1)\right) + I_{pi}(i_1) \tag{19}$$

The signal produced by the light scattered inside the camera body I_{pi} is the same for the two neighboring pixels, because the scattered light signal has a slow spatial variation. The signal produced by the directly reflected light and the scattered light in the scene (outside the camera box) in the white region of the tag is k_t times greater than in the black region, so that:

$$I_d(i_2) + I_{po}(i_2) = k_t\left(I_d(i_1) + I_{po}(i_1)\right) \tag{20}$$

The equation system (14), (15) can be solved and we have the solutions:

$$I_{pi}(i_1) = \frac{1}{k_t - 1}\left(k_t \cdot I_m(i_1) - I_m(i_2)\right) \tag{21}$$

$$I_d(i_1) + I_{po}(i_1) = \frac{1}{k_t - 1}\left(I_m(i_2) - I_m(i_1)\right) \tag{22}$$

If we use this method and that described in the previous paragraph then all the signals I_d, I_{pi} and I_{po} can be determined.

This last method is useful because it does not involve any hardware modifications of the ToF camera. The measured distance errors produced by the scattered light inside the camera body are generally greater than those produced by the scattered light in the scene. For measured distance errors correction it is not necessary to know the value of the contrast factor k_t , (see Fig. 3).

1.4 Example of error correction produced by the scattered light.

For the moment we did not have a structured light ToF camera, so we imagined a simple experimental setup to verify the method described above.

Fig. 4. A line shape region of the cylindrical vase surface (specular) reflects more light than the neighborhood region. At the top of the vase we placed a half white and half black sheet of paper.

The optical configuration of our experimental setup creates a small region with the two signals needed for the distance correction. The idea was to use a cylindrical object which placed in a proper position will produce a specular reflection of the light emitted by the camera. If this specular reflection is pointed exactly to the camera it will also appear in the image. The image of this specular reflection is a tiny short line like that in Fig. 4. The image pixel of the line receives more light than the neighboring pixels, so we have the two signals from two neighboring pixels. Unfortunately a little detail remains: we don't know the ratio of the light received by the two pixels, i.e. the "k" factor. To determine "k" we have added a calibration tag - which is a sheet of paper half white and half black. The actual value of factor "k" will give the best distance correction for both the white and black parts of the paper.

The experiment was performed in a space containing many different objects; in these conditions we have significant scattered light and a high amplitude perturbing signal.
The ratio of the light detected by a pixel witch receives the specular reflected light and a neighboring pixel is 6.024 for the best distance corrections, Fig. 5 (right panel).
In a usual case the contrast factor "k" has the highest value in the absence of any parasitic signal but in this case the signal is a vector and its value can be greater or lower depending on the phase of the parasitic signal.

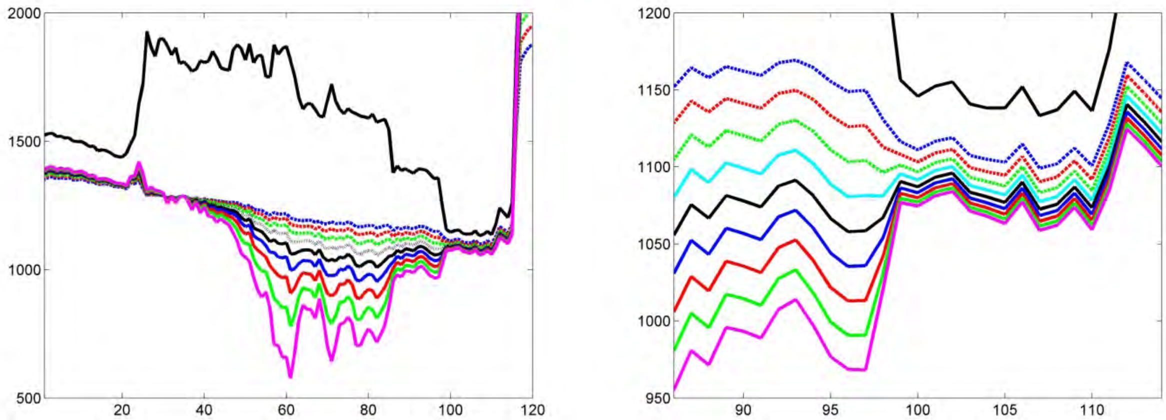

Fig. 5. Plot of the measured distance for a selected column of the distance image (left panel). A zoom of the plot in the region of the small tag half white (at the right) half black (at the left) is shown in the right panel. We observe that the white half is less affected by errors than the black half.

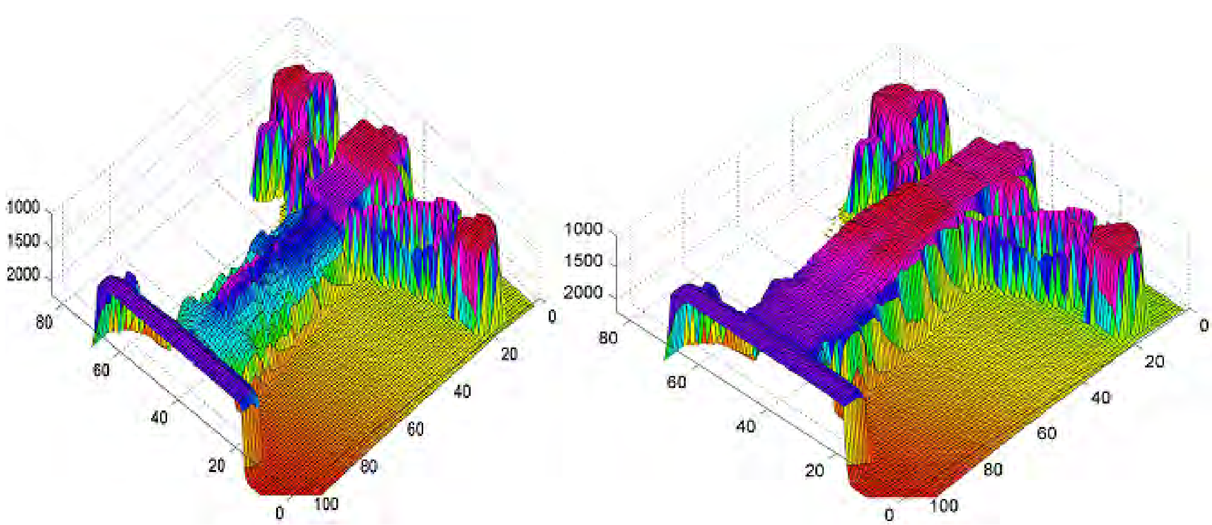

Fig. 6. 3D surf of the uncorrected image (left panel) and of the corrected image (right panel).

Obviously the correction will not be perfect but it is important to note that it is possible to improve the accuracy up to 10 times. The correction is dependent on the picture content.

As it can be observed in the pictures above the measured distance is dependent on the objects' reflectivity. This is an important disadvantage in industrial applications where the measured surface irregularities are in fact irregularities of the surface reflectivity.

1.5 The "fixed pattern" image error correction.

The amplitude and distance images from the ToF camera are affected by errors due to a "fixed pattern". These errors can not be avoided and black objects are affected the most. The fixed pattern image is the average image taken without any input light signal it is mainly produced by the dark current of the image sensor. Generally the signal from a pixel in the absence of any input light is a noise signal characterized by a mean value and a standard deviation. A "fixed pattern" image is an image containing the mean signal values for pixels. Because the dark noise signal of a light sensor depends on the temperature the fixed pattern image will be temperature dependent.

A fixed pattern image is obtained by averaging many "dark" images, for example with a covered objective. The fixed pattern image correction consists in the subtraction of this fixed pattern image from the image to be corrected.

The same procedure must be applied in the case of ToF cameras but we must note that the initial signal image is in fact a vector image with components I_x and I_y . If the fixed pattern image is the vector image with the components mean($I_{x0}(i)$) and mean($I_{y0}(i)$) then the fixed pattern ToF image correction consists in the vector subtraction of this fixed pattern image [$I_{x0}(i)$, $I_{y0}(i)$] from the vector image to be corrected [$I_{xm}(i)$, $I_{ym}(i)$].

$$I_x(i) = I_{xm}(i) - I_{xo}(i)\,,\quad I_y(i) = I_{ym}(i) - I_{yo}(i) \tag{23}$$

The corrected amplitude and distance images are computed with (7) and (8).

We point out that vector image processing is the best solution for the ToF cameras. Scalar (*non-vector*) algorithms for amplitude or distance image processing gives good results in most situations but not always as in the case of "fixed pattern" image errors correction.

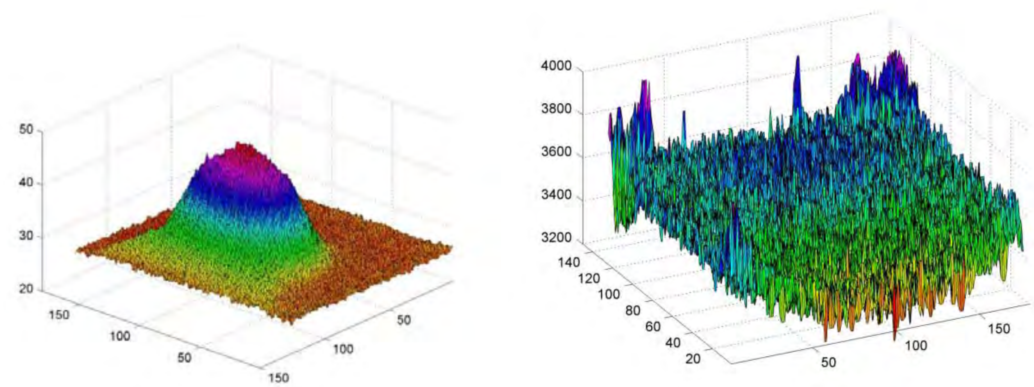

Fig.7. The fixed pattern amplitude image (left panel) and the fixed pattern distance image obtained by averaging 100 vector images (right panel).

The non-uniformities of the fixed pattern amplitude image are produced by non-uniform chip heating, Fig.7 (left panel). Because the dark distance image is noisier than the amplitude image averaging 100 vector images is not enough for the fixed pattern distance image, it is a plane surface situated at a distance equal to $\lambda/4$.

The signals [I_{x0} , I_{y0}] are obtained from the samples $S_{01},\ldots, S_{04}$. In the absence of any input light, these signals are random variables with a mean value $\mu(S_{0i}) = S_{\mu i}$ and their standard deviation $\sigma(S_{0i}) = S_{\sigma i}$ so that:

$$\mu\left(I_{x0}\right) = S_{\mu 1} - S_{\mu 3} \cong 0\,,\quad \sigma\left(I_{x0}\right) \cong \sqrt{2}\cdot S_{\sigma i}\,,\quad \mu\left(I_{y0}\right) = S_{\mu 2} - S_{\mu 4} \cong 0\,,\quad \sigma\left(I_{y0}\right) \cong \sqrt{2}\cdot S_{\sigma i}$$

We suppose that for all i $S_{\mu i} = S_{\mu}$, $\sigma(S_{\mu i}) \cong S_{\sigma}$, $\sigma(I_x) = \sigma(I_y)$. In this case we have:

$$\mu(a_0) = c_d \sqrt{I_{x\mu}^2 + I_{y\mu}^2 + \sigma^2(I_{x0}) + \sigma^2(I_{y0})} \cong \sqrt{2} \cdot c_d \cdot \sigma(I_{x0})$$

The dark phase signal is a random signal with a uniform distribution in the interval [0, 2π], and a mean value of π.

$$\varphi = \pi + \arctan\left(\frac{I_{y0}}{I_{x0}}\right)$$

Small technological imperfections and non-uniform heeting of the chip produces slightly different fixed pattern images.

The method presented corrects the measured distance errors due to the fixed pattern image if this image is known. The method is useful to correct the distance images of the actual ToF cameras and will help the designer improve these devices. The method can be useful to other distance measuring devices using the same (ToF) principle.

1.6 Example of an image affected by fixed pattern errors and its correction.

The test image is a small sheet of paper 5x10 cm half black and half white suspended by a string at approximately 1 m in front of the camera. The black part of the paper is only 90% black and the background of the scene is dark black. We took pictures of this simple scene with 4 exposure times from 0.2 ms to 1.6 ms in steps of powers of 2.

To represent graphically the measured distance to the white and black parts of the paper and also the background distance we have selected a representative column of the distance image. The measured distance along the pixels of this column is plotted in Fig. 8 (left panel).

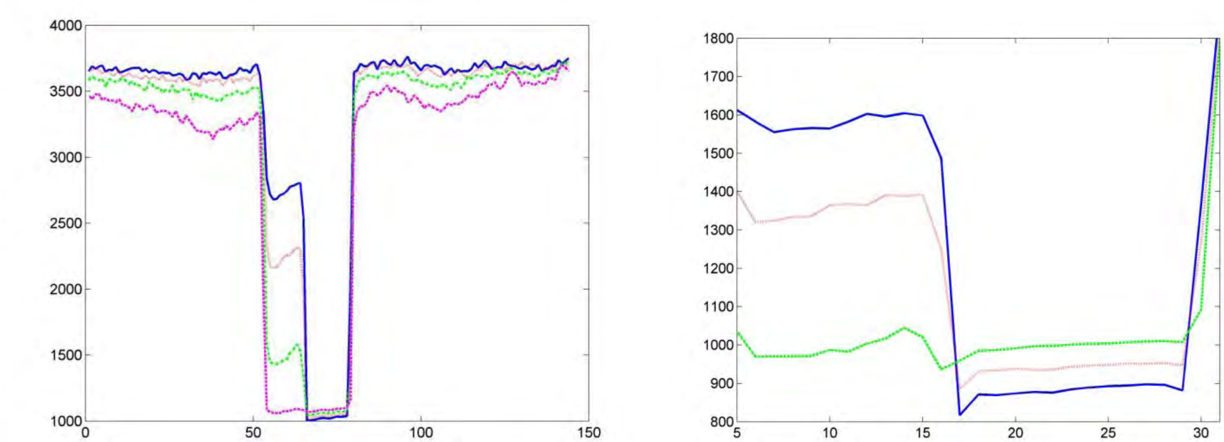

Fig. 8. The measured distance plot (left) and the corrected distance (right).

For the first three exposure times the measured distance to the black part of the paper is affected by large errors (1.7m for the 0.2 ms exp. time). For the 1.6 ms exp. time the measured distance to both parts of the paper is the same. We conclude that the fixed pattern correction is needed only for black objects and for low exposure times. That is what we expected because the fixed pattern vector $[I_{x0}, I_{y0}]$ affects the images only if its values are comparable with the values of $[I_x, I_y]$. In the regions where the image brightness is high the fixed pattern correction is comparable with the camera accuracy. The corrected images in these regions differ by less than 1%.

Another important thing is that the measured distance to the background is affected by errors in all cases. The dark black background is not placed exactly at 3.7 m and the measured distance to it changes with exposure time. Because the background is so dark the camera can not measure the distance to it, and the measured values are in fact the fixed pattern distance image so this behavior is quite normal. The distance to the background, in fact the fixed pattern distance, varies with the exposure time due to the light scattered inside the camera body.

To perform the distance image correction we need the fixed pattern amplitude and the distances images. So we removed the sheet of paper and we took the pictures of the background with the same exposure times as those used previously. A similar image can be obtained by covering the objective. However to obtain a useful fixed pattern image we had to average many identical images (1200).

We have applied the fixed pattern correction according to the method explained above. The corrected distance plot in the region of the black and white part of the paper is presented in Fig. 8 (right panel).

From this plot, it is clear that the measured distances are perfectly corrected for the 0.8 ms exposure time and the error is substantially reduced in the other two cases. In this example the fixed pattern is not the only factor which produces distance errors and the distance image needs to be corrected for the errors produced by scattered light.

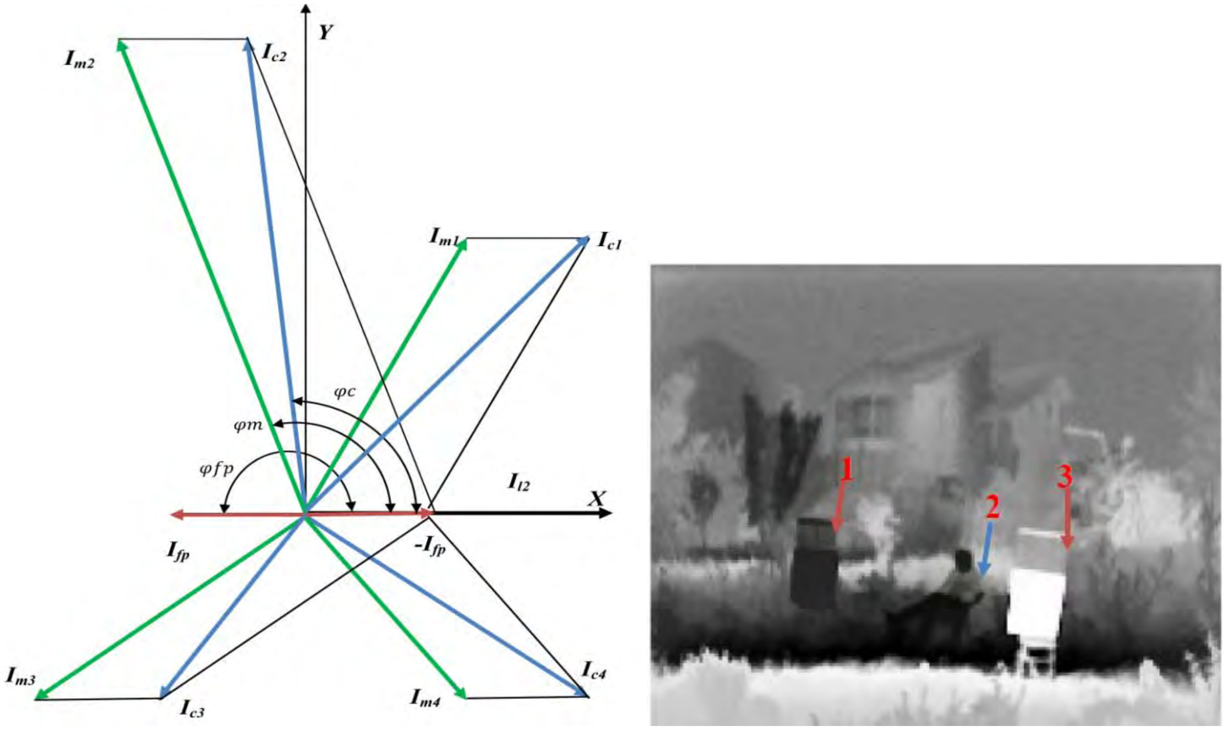

Fig. 9. The vector diagram of the measured and fixed pattern signals (left panel). A scene with different distance errors produced by the fixed pattern (right panel).

Let us point out some particular situations of the fixed pattern corrections. Some cases of distance error produced by the fixed pattern are graphically represented in Fig. 9. In the majority of cases the fixed pattern distance is half of the maximum no-ambiguity distance, so in the Fig. 9 represents this case.

If the distance to the object is lower than the fixed pattern distance the error introduced in this case is positive, and the measured distance is longer than the actual distance. The situation is reversed if the distance to the object is greater than the fixed pattern distance: in this case the error introduced is negative, and the measured distance is shorter than the actual distance.

In the distance image shown in Fig. 9 (right panel) the measured distance to the region marked 1 seems to be longer distance than that for the rest of the panel (waiter means a longer distance). The multicolor shirt marked 2 on the figure also appears to be at a longer distance than the rest of the body. The measured distance to the black part of the panel is shorter than the distance to the white part (darker means a shorter distance).

1.7 Conclusions and Acknowledgements

In our opinion ToF cameras will evolve towards a cheap and efficient device for many useful real time applications of image processing/analysis; one of the necessary developments is the elimination of a few sources of important errors or their automatic correction. Our model proved to be effective in the drastical reduction of the errors due to scattered light and we believe it is also effective in the reduction of other errors.

Our research was conducted 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) within the Information Society Technologies (IST) priority of the 6th Framework Program. This publication only 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.

[1] T. Oggier, M. Lehmann, R. Kaufmannn, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, N. Blanc, “An all-solid-state optical range camera for 3D-real-time imaging with sub-centimeter depth-resolution (SwissRanger)”, Proc. SPIE Vol. 5249, pp. 634-545, 2003.

[2] S. Oprisescu, D. Falie, Mihai Ciuc, Vasile Buzuloiu, “Measurements with ToF Cameras and their necessary corrections”, ISSCS, 2007, Iasi.

[3] SwissRanger SR-3000 Manual, Version 1.02 October, 2006.