비전 센서의 앨리어싱 방지 필터링 모방 기법

Emulation of Anti-alias Filtering in Vision Based Motion Mmeasurement

김 정 현[1]

Kim Jung Hyun[1]

Abstract This paper presents a method, Exposure Controlled Temporal Filtering (ECF), applied to visual motion tracking, that can cancel the temporal aliasing of periodic vibrations of cameras and fluctuations in illumination through the control of exposure time. We first present a theoretical analysis of the exposure induced image time integration process and how it samples sensor impingent light that is periodically fluctuating. Based on this analysis we develop a simple method to cancel high frequency vibrations that are temporally aliased onto sampled image sequences and thus to subsequent motion tracking measurements. Simulations and experiments using the 'Center of Gravity' and Normalized Cross-Correlation motion tracking methods were performed on a microscopic motion tracking system to validate the analytical predictions

Keywords: motion tracking, anti-aliasing, image stabilization, image vibration, motion estimation, motion filtering, video stabilization, visual servoing

1. Introduction

Capturing a digital image time sequence is a process that requires sampling of the space-time video volume both spatially and temporally, where each pixel integrates the light over time and space (Fig.1). Sampling is always accompanied by aliasing when the spatial and/or temporal signal content varies in frequencies that exceed the Nyquist frequency[1]. In this paper, we are concerned with temporal aliasing and especially its effect on visual motion tracking [9],[10]. In motion tracking, each temporally sampled image is processed using a variety of algorithms[2]-[4] to estimate the camera-object relative motion. When there is motion vibration and/or fluctuation in illumination, the photonic energy impingent on the image sensor will fluctuate. If this fluctuation contains frequency components exceeding the Nyquist frequency, its false low frequency 'alias' will appear in the temporally sampled image sequence, and thus in the motion tracking results.

A common practice in sampling of a continuous-time signal is to filter the signal before it is passed to the sampler. The filter used for this purpose is an analog low-pass filter whose cutoff frequency is not larger than the Nyquist frequency, called anti-aliasing filter[1]. However, in the case of video-time sampling, the physical signal

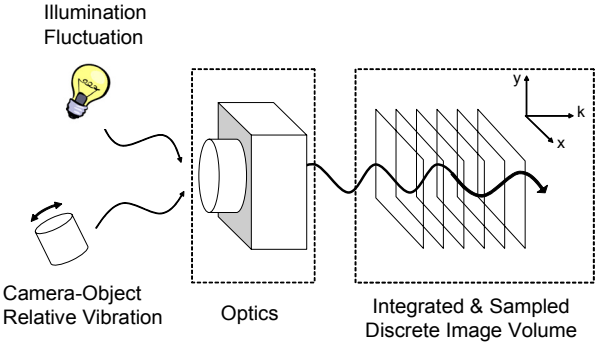

Fig. 1. Process of sampling of the space-time video volume.

Received : Aug.15.2010; Reviewed : Dec.14.2010; Accepted : Dec.30.2010
※ This research was supported by Kyungsung University Research Grants in 2010

[1] 경성대학교 메카트로닉스공학과 전임강사

medium is not an electrical voltage but photons impingent on a solid state array and building an anti-aliasing analog filter is not physically possible.

Current solutions to this problem, in the case of motion vibration, involve mechanically dampening the vibration either passively or using gyroscopes to measure and counter the vibration with opposing motion using servo motors; a field called image stabilization[5]-[7]. This approach is known to successfully eliminate artifacts, e.g. blurring, from high frequency motion up to the bandwidth of the servo motor, with sensitivity down to the resolution of the servo mechanism.

In this paper, we develop a novel method, Exposure Controlled Temporal Filtering (ECF) that utilizes the time-integration property of the image-exposure mechanism inherent in all imaging systems, to electronically cancel temporally aliased image intensity fluctuations caused by periodic motion vibration and/or illumination fluctuation without additional mechanical components.

We first present a theoretical analysis of the exposure induced image time integration process and how it samples sensor impingent light that is periodically fluctuating. The criterion for the cancellation of temporal aliasing is derived from this relationship. We also show analytically that temporal aliasing appears in the estimated motion parameters for motion tracking using first order image moments (FOIM)[2],[4], also known as the Center of Gravity method. The cancellation criterion also applies to this motion measurement. Simulation and experiment results are shown for motion tracking with FOIM and with normalized cross-correlation (NCC)[3].

2. Motion Tracking

Fig. 2 describes the process of motion tracking. Object-Camera relative motion will be projected onto the image plane through the optics, changing the impingent intensity $i(x, y, t)$. This continuous function is integrated and sampled to form the discrete function $I(x', y', k)$ $I'(x', y', k)$. At this stage the signal is physically an electronic

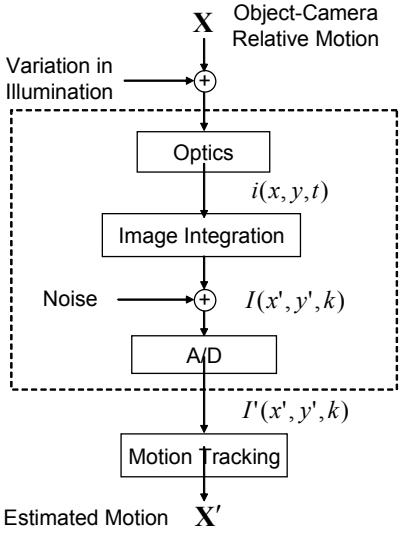

Fig. 2. Process of motion tracking

charge ready to be digitized. Thus the signal is subject to noise sources mainly of slowly varying thermal nature. The signal is then quantized to get $I'(x', y', k)$, a fully digital video volume. This function is then subject to a motion tracking process.

High frequency object-camera relative vibration will cause $i(x, y, t)$ to fluctuate at the same frequency. We are interested in how this fluctuation will affect the motion estimation result from motion tracking. The motion tracking method we choose for purpose of analysis is the widely used 'Center of Gravity' (COG) derived from first order Image moments[2]. An analytical model is derived and based on this, a strategy to cancel high frequency aliased measurements is presented.

2.1. COG of Miniscule Are

We start with defining a miniscule square area, $\triangle_x \times \triangle_y$ (Fig.3) on the image plane that has uniform intensity, ρ. The rest of the image plane is considered to have zero intensity. We assume that this area undergoes temporally continuous motions in both the x and y directions defined by $(X(t), Y(t))$. The generated continuous image $i(x, y, t)$ is integrated and sampled in time to obtain $I(x, y, k)$. The COG in the y direction obtained at time sample k is obtained from the equation,

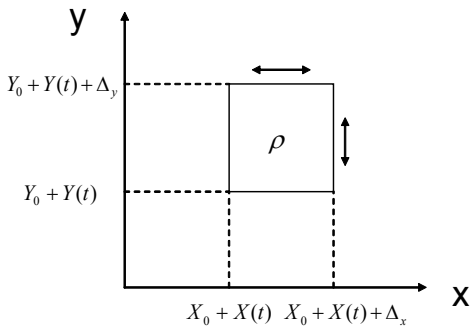

Fig. 3. Miniscule area with uniform intensity moves in x and y directions.

$$\bar{I} \cdot \bar{y}(k) = \iint_{XY} [y \cdot I(x, y, k)] \cdot dxdy \tag{1}$$

where on the right hand side is the definition of first-order image moment, $\bar{I}$ is the total intensity and $\bar{y}(k)$ denotes the COG y coordinate at sample k. All derivations are the same for x thus are omitted. The discrete time image function $I(x, y, k)$ is a time integration of $i(x, y, t)$, therefore (1) becomes,

$$\bar{I} \cdot \bar{y}(k) = \iint_{XY} y \cdot \int_{kT_s}^{kT_s+T_e} i(x, y, t)dt \cdot dxdy \tag{2}$$

Switch integration order and reflect the motion of the square onto the area integration range,

$$\bar{I} \cdot \bar{y}(k) = \int_{kT_s}^{kT_s+T_e} \int_{Y_0+Y(t)}^{Y_0+Y(t)+\Delta_y} \int_{X_0+X(t)}^{X_0+X(t)+\Delta_x} y\rho \cdot dxdydt \tag{3}$$

Solve and divide by $\bar{I} = \rho \cdot \triangle_x \triangle_y \cdot T_e$ to get,

$$\bar{y}(k) = Y_0 + \frac{1}{2}\Delta_y + \frac{1}{T_e} \int_{kT_s}^{kT_s+T_e} Y(t)dt \tag{4}$$

The result shows that $\bar{y}(k)$ is a time average of the continuous time function $Y(t)$ during the period of exposure. Further, if you transform the time integration into a convolution,

$$\bar{y}(k) = Y_0 + \frac{1}{2}\Delta_y + \frac{1}{T_e}\left[Y(t) * \prod_0^{T_e} \right]\Big|_{t=kT_s} \tag{5}$$

The convolution operation is a multiplication in the frequency domain. The amplitude of the Fourier Transform of the square function $\Pi_0^{T_e}$ is

$$\left\| \Im\left(\Pi_0^{T_e}\right) \right\| = T_e \cdot \sin c\left(\frac{T_e\omega}{2}\right) \tag{6}$$

a sync function which goes to zero at points where the condition $T_e\omega/2 = n \bullet \pi (n = 0, \pm 1, \cdots)$ is satisfied. When there is periodic motion in $Y(t)$ with the period T_m i.e. when $\omega = 2\pi / T_m$, due to the convolution effect, those motions will be cancelled when the ratio between exposure time T_e and function period T_m are an integer, i.e.

$$\frac{T_e}{T_m} = N, \quad N = 1, 2, 3, \ldots \tag{7}$$

is satisfied. The motion $Y(t)$ can be considered as being filtered by a filter having a sync function as its frequency response. This means that if the continuous motion $Y(t)$ contains periodic motion vibrations, the effect of these vibrations can be cancelled in $\bar{y}(k)$ if the exposure time is appropriately tuned such that cancellation criterion (7) is satisfied.

2.1.1 COG of General Object

Now we consider a general rigid body object projected on the image plane (Fig. 4) consisting of an infinite number of minuscule square areas (Fig.3) with different intensity values. When the object undergoes motion, all minuscule areas go through identical motion since the object is considered rigid. We define coordinate frame {C} (Fig.4) as a moving one attached to the object at its COG. {O} is a global reference coordinate frame which is stationary.

We start with considering each miniscule square of the rigid object as having its own COG (y-axis), i.e.,

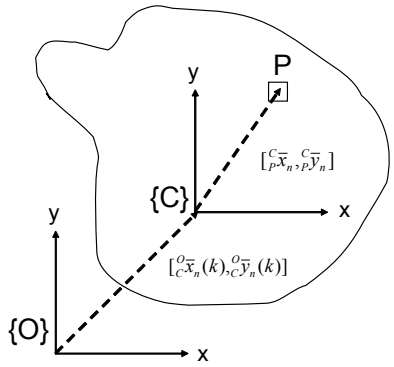

Fig. 4. Image of general object consisting of infinite number minuscule areas. {O} denotes the global coordinate frame; {C} denotes the coordinate frame attached to the moving object at its COG.

$$\bar{I}_n \bar{y}_n(k) = \int_X \int_Y y \cdot \left[\int_{kT_s}^{kT_s+T_e} {}^{n}i(x,y,t) \cdot dt \right] \cdot dydx \quad (n = 1,2,...\infty) \tag{8}$$

Summing for the whole object area we get,

$$\sum_{n=1}^{\infty} \bar{I}_n \bar{y}_n(k) = \int_X \int_Y y \cdot \left[\int_{kT_s}^{kT_s+T_e} \sum_{n=1}^{\infty} {}^{n}i(x,y,t) \cdot dt \right] \cdot dydx \tag{9}$$

However the COG y-axis component for the whole area, $\bar{\bar{y}}(k)$, by definition is,

$$\bar{\bar{I}} \cdot \bar{\bar{y}}(k) = \int_X \int_Y y \cdot \left[\int_{kT_s}^{kT_s+T_e} {}^{n}i(x,y,t) \cdot dt \right] \cdot dydx \tag{10}$$

The right hand side of (9) and (10) are identical therefore,

$$\bar{\bar{I}} \cdot \bar{\bar{y}}(k) = \sum_{n=1}^{\infty} \bar{I}_n \bar{y}_n(k) \tag{11}$$

Since we assume that the object undergoes rigid body motion the COG of minuscule area P (Fig.4) can be presented as,

$$\bar{y}_n = {}^{O}_{P}\bar{y}_n(k) = {}^{O}_{C}\bar{y}(k) + {}^{C}_{P}\bar{y}_n \tag{12}$$

Note that ${}^{O}_{C}\bar{y}(k)$ is common for all n, thus the subscript is omitted, and ${}^{C}_{P}\bar{y}_n$ is time invariant for all n. Substitute (12) into (11),

$$\bar{\bar{I}} \cdot \bar{\bar{y}}(k) = \left[\sum_{n=1}^{\infty} \bar{I}_n \right] \cdot {}^{O}_{C}\bar{y}(k) + \sum_{n=1}^{\infty} \bar{I}_n \cdot {}^{C}_{P}\bar{y}_n \tag{13}$$

Since the origin of {C} is attached at the object COG, the second term on the right hand side of (13) is zero and since,

$$\bar{\bar{I}} = \sum_{n=1}^{\infty} \bar{I}_n \tag{14}$$

we obtain,

$$\bar{\bar{y}}(k) = {}^{O}_{C}\bar{y}(k) \tag{15}$$

The significance of (15) is that it shows the COG obtained from the whole object is equal to the time varying portion, ${}^{O}_{C}\bar{y}(k)$, of the COG for each minuscule area (Fig.5). And the behavior of the COG of a minuscule area

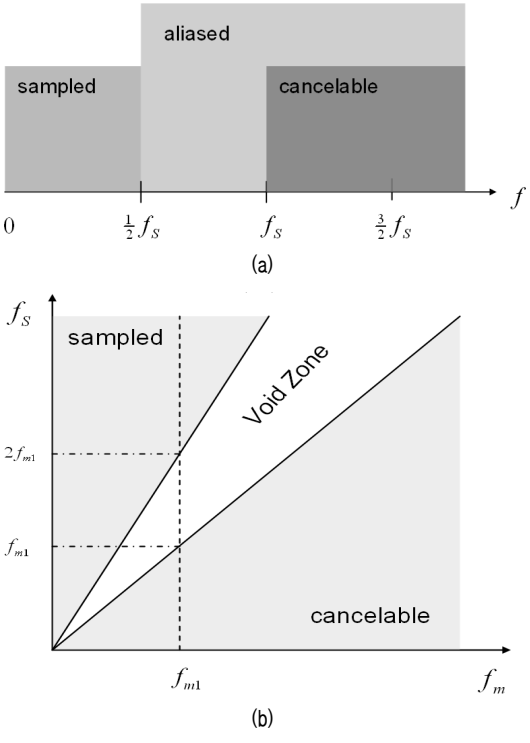

Fig. 5. (A) Plot showing regions of sampling, aliasing and exposure canceling. (B) Dependency of *Sampled* and *cancelable* regions on sampling frequency *fs*.

is such that if the exposure time is appropriately tuned, periodic motion vibrations that have periods which satisfy the cancellation criterion (7) will not show up in the measurement. Thus, the cancellation criterion also applies to a general object in the image plane.

2.1.2 Cancellation of Aliased Signal

The primary objective of ECF is to cancel the aliasing of high frequency fluctuations which are either from motional vibrations or illumination fluctuations. We have shown that motion tracking measurements of periodic motions will be cancelled when those motion periods and exposure time satisfy cancellation criterion (7). This is done by tuning the exposure time T_e such that (7) is satisfied for a known fluctuation period T_m. However, there exists a physical limit on the range of achievable exposure times T_e, which is $T_e \leq T_s$, where T_s is the sampling period. This limits the frequency range of the signals that ECF can cancel, i.e. only when $f_m \geq f_s$ can T_e be tuned such that the cancellation criterion is satisfied.

To explain more clearly, Fig.5-(A) shows three distinct regions in the frequency domain: 1) when the fluctuation frequency f_m is below the Nyquist frequency $f_s/2$, the signal can be *sampled* without aliasing; 2) when it is at or above the Nyquist frequency, the signal will cause hidden oscillations or will be aliased, both of which are unwanted; and 3) when it is at or above the sampling frequency f_s, it can be cancelled through ECF. The problem arises when $f_s/2 \leq f_m < f_s$ since the signal will be aliased and it cannot be cancelled. This range is from now on referred to as the *void zone*.

In order to resolve this problem, we choose a strategy of varying the sampling frequency f_s which can be done easily in many modern cameras. The plot in Fig. 5-(B) shows three regions; *sampled, cancelable* and *void zones* and how the range of these regions changes depending on the sampling rate f_s. The strategy is to change both the sampling period T_s and the exposure time T_e. For example, when the frequency of the fluctuation is f_{m1} (Fig. 5-(B)), choose f_s larger than $2f_{m1}$ when sampling is intended. For cancellation, choose f_s smaller than f_{m1} then tune T_e such that criterion (4) is satisfied for cancellation.

3. Simulation Results

Simulations that verify the analytically derived results were performed. Object motion was simulated by a square on the image plane having uniform intensity undergoing sinusoidal vibration according to the specified frequency against a zero-intensity background (Fig. 8). This forms the impingent intensity function $i(x, y, t)$, In order to simulate its temporal and spatial continuousness, a 800x800 pixel video sequence with time resolution of 50 micro-seconds was used. We spatially and temporally integrate this function within each pixel area and exposure time respectively to obtain $I(x, y, k)$, which is a time sequence of 40x40 pixel images. The COG motion tracking algorithm is then applied to the image sequence.

Fig. 6 shows the time plot and a sample image for three different T_e/T_m values. In the time plots, the highly varying

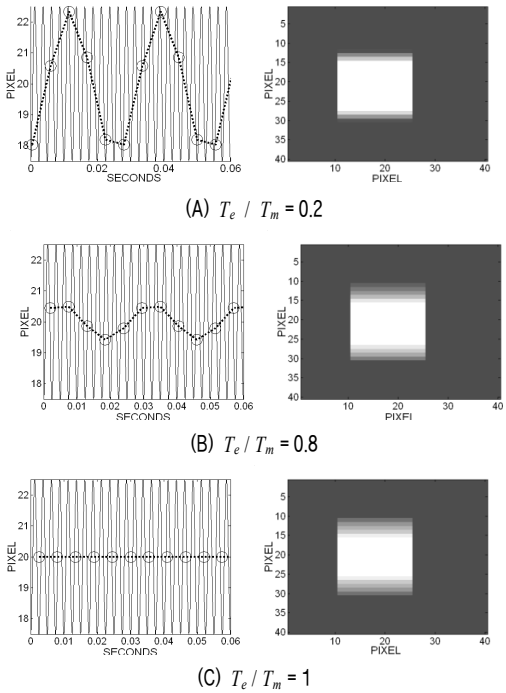

Fig. 6. Varying T_e with constant T_s = 5.5 msec and T_m = 2.5 msec. Left column is time plot of real motion versus motion tracking result and right column is a sample image at each respective exposure time.

plot is the simulated motion, and the circle-dotted line plots correspond to the estimation results from motion tracking. In each case, sampling time T_s is 5.5 milliseconds while the motion period T_m is 2.5 milliseconds, thus aliasing occurs in all three cases. However, the exposure time differs for each case, causing a change in estimated motion amplitude, and finally the sinusoidal motion is totally cancelled when $T_e / T_m = 1$ (Fig.6-(c)) as predicted.

Note that as the exposure time is increased, from $T_e / T_m = 0.2$ to 1.0, image blur increases. This is unwanted in most imaging situations where image clarity is critical. This does not however disrupt motion tracking with COG even when the motion is quite large as in the simulation. However, other motion tracking methods can be degraded with increasing blur, especially when motion range is larger. For example, results from Normalized Cross-Correlation (NCC), for which experimental results are shown in Sec. V, indicate that the peak correlation strength is weakened when the amplitude of the periodic vibration increases.

Fig. 7 compares the analytical model prediction with the simulation results. The exposure time was varied incrementally while the motion and sampling periods were kept constant. It is shown that the simulation results conform well to the analytical predictions.

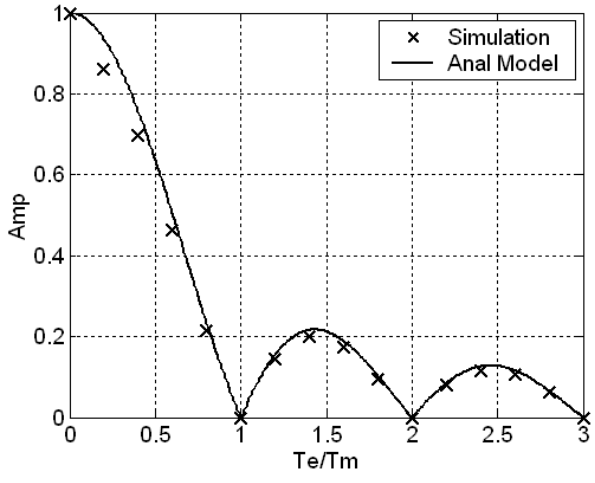

Fig. 7. Comparison of analytical and simulated Amplitude vs. T_e / T_m Plots.

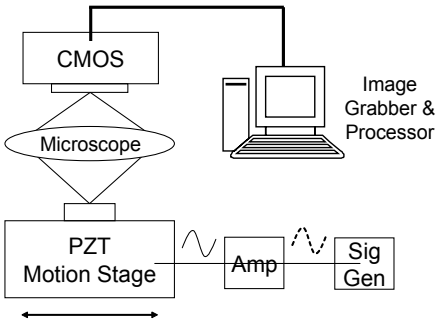

Fig. 8. Experimental Setup.

4. Experiment Results

Experiments were performed to verify the analytical predictions and the simulation results. Fig. 10 shows a diagram describing the experimental setup used. The periodic motion was generated by a piezo-electric motion stage (P.I. polytech) which was driven by a signal generator. A micro cantilever (Fig. 12) that was fastened to the stage was imaged using an optical microscope. A high-speed CMOS camera (Mikrotron) with variable sample and exposure times was attached to the microscope to capture images which were subsequently sent to the image grabber/processor PC unit for motion tracking..

Pure sinusoidal motion of 65 Hz was commanded into the motion stage by the signal generator. The motion-stage which carried the micro cantilever moved in the horizontal direction orthogonal to the optical imaging axis. The sample rate of the camera was set at 50 Hz, therefore the maximum possible exposure time was 1/50 sec. The exposure time of the CMOS camera was variable with a resolution that was sufficient for our purposes. The commanded 65 Hz periodic motion of the micro cantilever was measured while varying the exposure time. Fig. 9 shows the time plot (left) and frequency domain plot (right) of the measured motion using COG method, when the exposure time was at $T_e = 1 / 82$ sec and $T_e = 1 / 65$sec, respectively. In both cases the sampling rate was 50 Hz. When $T_e = 1 / 82$ the 65 Hz motion measurement is aliased to 15 Hz as can be clearly seen on Fig. 9-(A) (right). The small peak at 5 Hz represents the third harmonic of the motion vibration aliased onto the measurement, i.e. $4f_s$-$3f_m$

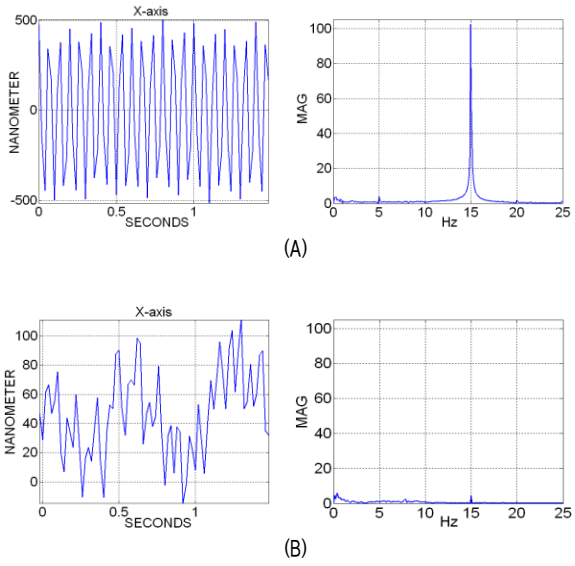

Fig. 9. Experimental COG Motion measurement time plot (left) and frequency domain plot (right) when T_m=1/65 sec, T_s = 1/50 sec and (A) Te=1/82 sec (T_e / T_m = 0.793), (B) T_e = 1/65 sec (T_e / T_m = 1).

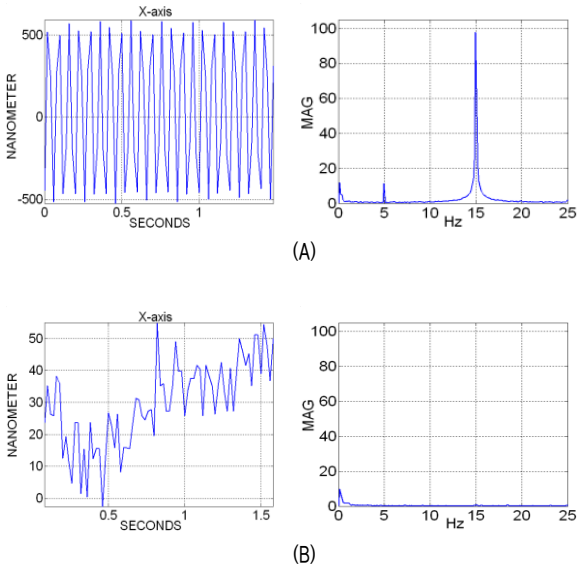

Fig. 10. Experimental NCC Motion measurement time plot (left) and frequency domain plot (right) when T_m = 1/65 sec, T_s = 1/50 sec and (A) T_e = 1/199 sec (T_e / T_m = 0.322), (B) T_e = 1/65 sec (T_e / T_m = 1).

= 5*Hz*. Apparently, the sinusoidal motion commanded by the signal generator is not totally pure and has higher harmonics. However, when $T_e = 1/65$ sec, criterion (7) was satisfied, i.e. T_e / T_m =1, and the aliased signal was practically cancelled (Fig.9-(B), right). The aliased third harmonic at 5*Hz* has also disappeared as is expected, since it corresponds to T_e / T_m =3, which also satisfies the cancellation criterion (7). There still remains oscillating components in the time plot of Fig.9-(B) that can be seen as the very small peaks in the frequency plot Fig.9-(B) (right). They are from electronic noise and from remnant 15 *Hz* oscillation. The remnant 15*Hz* peak is due to the limitation of accuracy with which the exposure time can be controlled. However, the impact of the remnant peak is negligible as its magnitude is comparable with other electronic noise sources. Finally, the slowing varying component remaining after the cancellation shown in the time plot of Fig.9-(B) is from thermally induced structural drift between the camera and object.

Fig. 10 shows a similar experiment for motion measurement, in this case using the Normalized Cross-Correlation (NCC) method[2]. NCC is a well known and widely used technique for motion tracking. Again, the 65 Hz motion was aliased to 15 Hz (Fig.10-(A), right), however was practically cancelled when T_e=1/65 sec (Fig.10-(B)) and criterion (4) was satisfied. Again the remnant peak in the frequency plot again is seen but with negligible effect. Also the slowly varying component in the time plot (Fig.10-(B)) accounts for thermal drift.

Fig. 11 shows that experimental results conform well to the analytical model. The exposure time was varied in twelve steps for both COG and NCC, and the measured amplitude was compared with the analytical prediction.

Fig. 12 shows images taken for a micro cantilever when the exposure times were at $T_e = 1/199$ sec and $T_e = 1/65$ sec, respectively. As was shown in the simulations, when the exposure time is increased the image of the object is more blurred. The image on the left is clearer especially on the edges compared to the one the right. In the case of COG, image blur does not disrupt motion tracking even when the motion is quite large. However, in the NCC method, object position is determined by finding the peak of a correlation function between a predefined model image of the object and the target image presumably

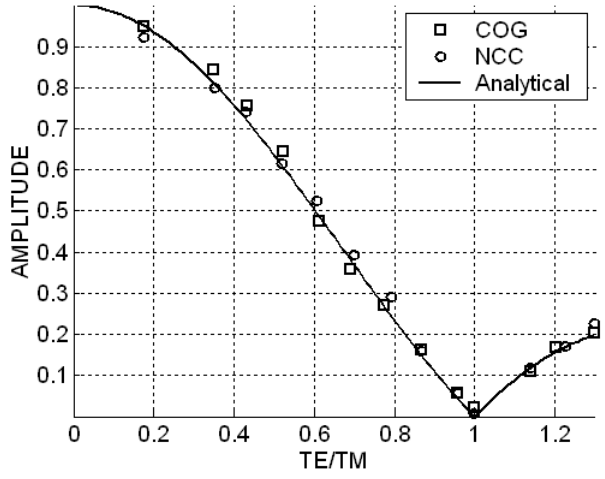

Fig. 11. Comparison between Analytical prediction and experimental results for motion measurements from COG and NCC methods.

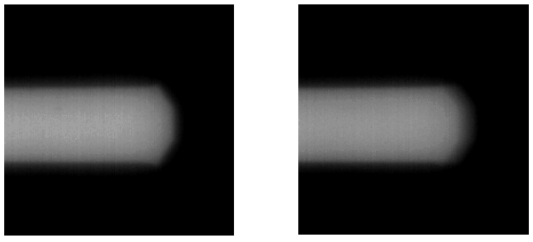

Fig. 12. Images of a micro cantilever vibrating at 65 Hz along the horizontal y-direction. The camera sample rate is 50 Hz and exposure times are $T_e = 1/199$ sec (left) and $T_e = 1/65$ sec (right).

containing the object[3]. The object in the target image will increasingly blur with increasing exposure time and increasing motion amplitude, which will weaken the peak correlation strength. When dealing with relatively small structural vibrations this effect is not a concern, however further considerations are recommended for larger motions. Creating a *blurred* model image based on known motion periods is a possible solution.

5. Conclusions and Remarks

This paper presented a novel method ECF that can be applied to visual motion tracking and that can cancel the temporal aliasing of periodic vibrations and fluctuations in illumination through the control of exposure time. Based on a theoretical analysis of the mechanism of image exposure a cancellation criterion was established. Along with the cancellation criterion, a simple method was developed, in which the image exposure time is controlled to cancel temporally aliased motion tracking measurements induced by high frequency periodic vibrations and fluctuations. The method can be useful inside a visual servo framework functioning as an *anti-aliasing filter* for canceling the aliasing of high frequency oscillations. This method can be applied especially to microscopic imaging systems that suffer from structural vibrations, a common problem in such systems where mechanical solutions are either too expensive or cannot achieve the required precision. Any other vision systems where conventional optical image stabilization techniques are either not feasible or too expensive are also good candidates for application.

참고문헌

[1] B. Porat, “A Course in Digital Signal Processing,” John Wiley & Sons, 1997.

[2] W.T. Freeman, D.B. Anderson, P.A. Beardsley, C.N. Dodge, M. Roth, C.D. Weissman, W.S. Yerazunis, W.S., Kage, H., Kyuma, K., Miyake, Y.; Tanaka, K., “Computer Vision for Interactive Computer Graphics”, *IEEE Computer Graphics and Applications*, Vol.18, Issue3, pp.42-53, May-June 1998

[3] L. Di Stefano, S. Mattoccia, M. Mola,” An efficient algorithm for exhaustive template matching based on normalized cross correlation,” Image Analysis and Processing, 2003. Proceedings. 12th International Conference on 17-19 Sept. 2003 pp. 322-327

[4] B.K.P Horn, “Robot Vision”, The MIT Press, Cambridge, MA, 1986

[5] R. Kurazume, S. Hirose, “Development of image stabilization system for a remote operation of walking robots,” Proc. IEEE Int. Conf. on Robotics and Automation, pp.1856-1861, 2000.

[6] J.S. Jin, Zhigang Zhu; Guangyou Xu, ”A stable vision system for moving vehicles,” Intelligent

Transportation Systems, IEEE Transactions on Vol.1, Issue1, March 2000 pp.32-39

[7] Canon: US Patent #6,606,456, "Image-shake correcting device" (Fujinaga; Nobuhiro, 2002).

[8] Mikrotron; http://www.mikrotron.de

[9] J. Kim, S.K. Kuo, C.H. Menq, "An Ultra Precision Six-Axis Visual Servo Control System," IEEE Transactions on Robotics, to appear in October 2005 issue.

[10] J. Kim, C. H. Menq, "Visually servoed 3D alignment of multiple objects with sub-nanometer precision," IEEE Trans. Nanotechnol., Vol.7, No.3, pp.321-330, May 2008.

김 정 현

1999 서울대학교 기계항공우주공학부(공학사)

2007 오하이오 주립대학교기계공학전공(공학박사)

2010~현재 경성대학교 메카트로닉스공학과 전임강사

관심분야 : Precision Motion Control, Machine Vision