Eccentricity Estimator for Wide-Angle Fovea Sensor by FMI Descriptor Approach

Sota SHIMIZU, *Member, IEEE* and J. W. BURDICK, *Member, IEEE*

Abstract— **This paper proposes a method for estimating eccentricity that corresponds to an incident angle to a fovea sensor. The proposed method applies Fourier-Mellin Invariant descriptor for estimating rotation, scale, and translation, by taking both geometrical distortion and non-uniform resolution of a space-variant image by the fovea sensor into account. The following 2 points are focused in this paper. One is to use multi-resolution images computed by Discrete Wavelet Transform for reducing noise caused by foveation properly. Another is to use a variable window function (although the window function is generally used for reducing DFT leakage caused by both ends of a signal.) for changing an effective field of view (FOV) in order not to sacrifice high accuracy. The simulation compares the root mean square (RMS) of the foveation noise between uniform and non-uniform resolutions, when a resolution level and a FOV level are changed, respectively. Experimental results show that the proposed method is consistent with the wide-angle space-variant image by the fovea sensor, i.e., it does not sacrifice high accuracy in the central FOV.**

I. INTRODUCTION

LOG-POLAR mapping, biologically-inspired by analytic formulation of cortical mapping of the primate visual system [1], is well-known as a method that does not only reduces image data size drastically but also gives rotation- and scale-invariant (RS-invariant) property. In order to acquire the log-polar image, not only by software methods [2] but also by 2 kinds of hardware methods have been proposed and developed. One is a method by a special-made CCD/CMOS chip, i.e., a retina-like sensor, where united size of photosensitive elements increases logarithmically as going to periphery [3]-[5]. The other is a method by special-made optical lens where projected image is highly distorted geometrically, i.e., a wide-angle foveated (WAF) lens [6]-[8]. In the latter case, the special-made lens is usually combined with a commercially-available Cartesian (i.e., linear-coordinate) vision chip, i.e., where the size of photosensitive elements is uniform (on the other hand, the former is combined with a commercially-available conventional lens). The authors have been doing research on fovea vision in terms of foveation modeling and its applications by using an actual WAF lens [9]-[11]. It is

well-known that a conventional Cartesian vision sensor acquires an input image with uniform-resolution and gives it translation-invariant property. However, the resolution by the fovea sensors is not uniform in the entire field of view (FOV), it is the highest in the central FOV, and it decreases rapidly as going to periphery. That is, the input image by the fovea sensor has no longer a translation-invariant property in Discrete Fourier Transform (DFT), even if it is transformed into linear coordinates. On the other hand, as introduced in the opening sentence, the fovea sensor has an advantage that it can give us the log-polar image that is RS-invariant around the optical axis of the fovea sensor.

Figure 1(a) shows a linear-coordinate Target image *I* of 512x512 [pixels]. Figures 1(b) and (c) are a distorted foveated image *DF*, computed from the image *I* by Advanced Wide-Angle Foveated (AdWAF) model [8], and an undistorted foveated image *UDF*, remapped from the image *DF* into a linear-coordinate space, in 3 cases of eccentricity, that corresponds to an incident angle to the sensor (both images have 128x128 [pixels] as their size). Watermark-like arias seen in the image *UDF* of Fig.1(c) increases as the eccentricity gets larger. Such alias occurs when the image is up-sampled from low resolution to high resolution. Figure 1(d) is the original log-polar image *OLP* that is transformed from a part of the image *DF* into polar coordinates. In order to keep the RS-invariant property mentioned above, we need to compensate deformation in the image *OLP* (this deformation increases highly as the eccentricity gets larger). Figure 1(e) is the compensated log-polar image *CLP* where the deformation is corrected from the image *OLP* by the eccentricity compensator (EC) [9]. It is noted that the image quality decreases as the eccentricity gets larger.

A log-polar mapping of DFT magnitude of an input image is well-known as Fourier-Mellin Invariant (FMI) descriptor [12]-[14]. The input image by the fovea sensor is space-variant, originally. This means there are 2 factors, i.e., one is geometrical distortion of the image and the other is non-uniform resolution. Thus, we need to take the non-uniform resolution into account in order to apply FMI descriptor for the space-variant image, even if the geometrical distortion is removed. Section II of this paper proposes eccentricity estimator (EE) that estimates the eccentricity by applying FMI descriptor for multi-resolution images by Discrete Wavelet Transform. In addition, a variable window function is introduced in order not to sacrifice high accuracy in the central FOV. Section III examines the proposed method by a simulation of foveation noise from white Gaussian noise images and experiments.
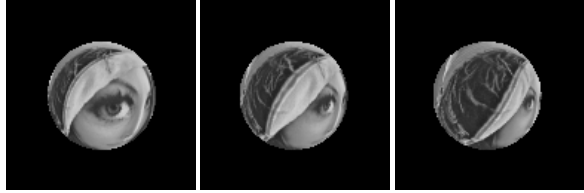
(a)Linear-coordainate target image (*I*) (of 512x512 [pixels])

(b)distorted foveated image (*DF*) (of 128x128)
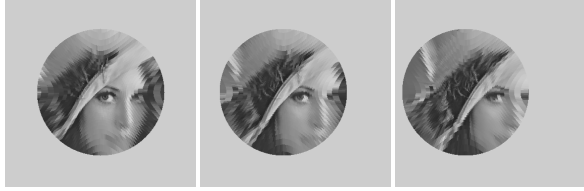
(c)undistorted foveated image (*UDF*) (of 128x128)

(d)original log-polar image (*OLP*) (of 128x64)

(e)compensated log-polar image (*CLP*) (of 128x64)

Fig. 1 Target image *I*, distorted foveated image (*DF*), undistorted foveated image (*UDF*), original log-polar image (*OLP*) and compensated log-polar image (*CLP*) in each eccentricity θ_ε (=0, 18.69 and 34.08[°] from the left)

II. ECCENTRICITY ESTIMATOR

A. Linear-coordinate Image with Space-variant Resolution

As shown in Fig. 1(b), an input image acquired from the fovea sensor is distorted highly because the image height from the image center (corresponding to its optical axis) is not linear to the object height. This paper applies Advanced Wide Angle Foveated (AdWAF) model [8] for a concrete foveation model of the fovea sensor. The AdWAF model combines planar and spherical projections with linear coordinates and logarithmic coordinates, respectively. Thus, the FOV is divided into 4 regions: fovea ($0 \leq \theta \leq \theta_0$),

para-fovea ($\theta_0 \leq \theta \leq \theta_1$), near-periphery ($\theta_1 \leq \theta \leq \theta_2$), and periphery ($\theta_2 \leq \theta \leq \theta_{\max}$). Figure 2 compares an AdWAF image and a linear-coordinate image by pinhole camera (PHC) lens model. In this simulation, image intensity is changed in order to make it easy to distinguish boundaries of the FOV, i.e., θ_0= 9.826[°], θ_1=19.107[°], θ_2= 34.715[°], and $\theta_{\max}$= 60.0[°]. The AdWAF model and the PHC lens model are defined in the following.

AdWAF model:

$$r = \begin{cases} r_{\max} c_0 f_1 \tan\theta & (0 \leq \theta \leq \theta_0) \\ r_{\max}\{c_1 \log_a f_1 \tan\theta + d_1\} & (\theta_0 \leq \theta \leq \theta_1) \\ r_{\max}\{c_2 \log_b (f_2\theta) + d_2\} & (\theta_1 \leq \theta \leq \theta_2) \\ r_{\max}\{c_3 f_2\theta + d_3\} & (\theta_2 \leq \theta \leq \theta_{\max}) \end{cases} , \quad (1)$$

where f_1 and f_2 are focal lengths for planar projection and spherical projection, basis *a* and basis *b* are given by $a = \exp(1/f_1 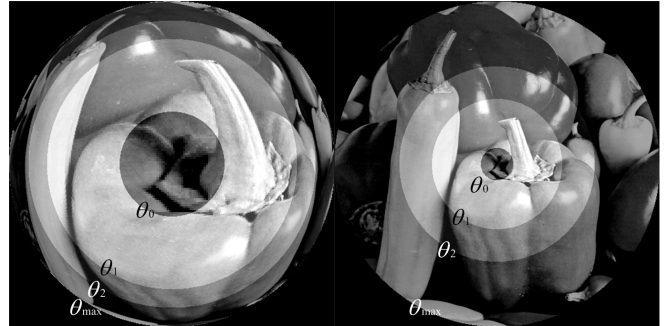\tan\theta_0)$ and by $b = \exp(1/f_2\theta_2)$, respectively. The $r_{\max}$ is the maximum image height when $\theta=\theta_{\max}$, c_i (*i*=0, 1, 2, 3) is a scale modification factor for adjusting image height partly in each region, and d_i (*i*=1,2,3) is determined from continuity of the image height and its magnification.

PHC lens model:

$$r = L \tan\theta \qquad\qquad (0 \leq \theta \leq \theta_{\max}) . \qquad (2)$$

where *L* is a distance from the optical center O_c to the object plane,

$$L = \frac{r_{\max}}{\tan\theta_{\max}} . \qquad (3)$$

(a)AdWAF image   (b)linear-coordinate image

Fig. 2 Comparison of AdWAF image and linear-coordinate image

Assume that a target image *I* is in the direction of eccentricity θ_ε and its rotation ϕ_ε on the object plane. That is, the eccentricity is an incident angle to the fovea sensor, as shown in Fig.3.

The incident direction, (θ, ϕ), from a point (*x, y*) on the image *I* to the optical center of the fovea sensor is expressed,

$$\begin{cases} \theta = \tan^{-1}\{\sqrt{(x + \varepsilon\cos\phi_\varepsilon)^2 + (y + \varepsilon\sin\phi_\varepsilon)^2}/L\} \\ \phi = \tan^{-1}\{(y + \varepsilon\sin\phi_\varepsilon)/(x + \varepsilon\cos\phi_\varepsilon)\} \end{cases} , \quad (4)$$

where ε is positional eccentricity on the object plane.

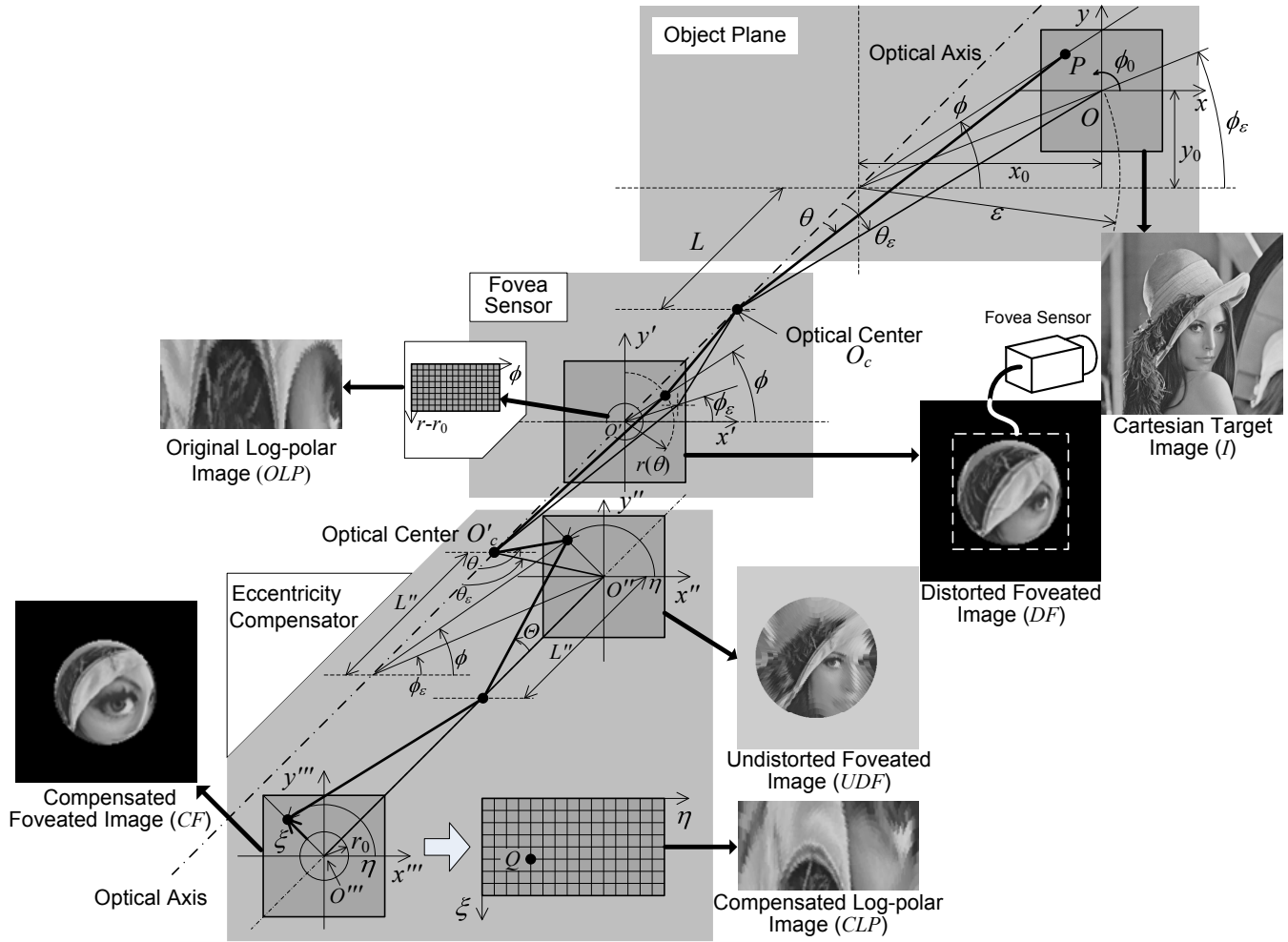$$\varepsilon = L \tan\theta_\varepsilon . \qquad (5)$$

Fig. 3 Coordinate systems of eccentricity estimator (EE) and eccentricity compensator (EC)

Coordinates (x', y'), of a distorted foveated image DF that corresponds to the input image by the fovea sensor, are given by eq. (6).

$$\begin{cases} x' = \alpha_1 r(\theta) \cos \phi \\ y' = \alpha_1 r(\theta) \sin \phi \end{cases} , \qquad (6)$$

where $r(\theta)$ is the image height of the fovea sensor in terns of incident angle θ, and α_1 is magnification of the image DF.

Coordinates (x'', y''), of an undistorted foveated image UDF with linear-coordinates, are expressed in eq. (7).

$$\begin{cases} x'' = \alpha_2(L \tan \theta \cos \phi - \varepsilon \cos \phi_\varepsilon) \\ y'' = \alpha_2(L \tan \theta \sin \phi - \varepsilon \sin \phi_\varepsilon) \end{cases} , \qquad (7)$$

where α_2 is magnification of the image UDF.

Both the image DF and the image UDF have space-variant resolution. Both images are mapped using mappings f_{DF} and f_{UDF} of an image warping technique, respectively.

$$f_{DF} : I \rightarrow DF , \qquad (8)$$

$$f_{UDF} : DF \rightarrow UDF , \qquad (9)$$

B. Fourier-Mellin Invariant descriptor

Figure 4(a) shows the image f and the image g transformed from the f with rotation ϕ_0 [rad], scale μ, translation (x_0, y_0).

$$\begin{aligned} g(x, y) = f(&\mu(x \cos \phi_0 + y \sin \phi_0) - x_0, \\ &\mu(-x \sin \phi_0 + y \cos \phi_0) - y_0) \end{aligned} . \qquad (10)$$

Fourier-Mellin Invariant (FMI) descriptor based on Fourier Mellin Transform is well-known for extracting rotation-, scale-, and translation-invariant feature [12]-[14]. FMI descriptor is equivalent with a log-polar mapping of DFT magnitude of the input image, around the DC component. This paper refers to this method for realistic implementation [13][14] in order to compute rotation, scale, and translation between the 2 images. FMI descriptor is derived simply in the following.

Denote DFTs of the image f and the image g as $F(m, n)$ and $G(m, n)$, and their log-polar mappings as f_{lp} and g_{lp}.

$$g_{lp}(\phi, \lambda) = \mu^{-2} f_{lp}(\phi - \phi_0, \lambda - \sigma) , \qquad (11)$$

where $\lambda = \log \sqrt{m^2 + n^2}$ and $\sigma = \log(\mu)$. Note that the log-polar mapping transforms rotation and scale to translation. DFTs of the log-polar mappings f_{lp} and g_{lp} are expressed as

$$G_{lp}(\upsilon,\varpi) = \mu^{-2}e^{-j2\pi(\upsilon\sigma+\varpi\phi_0)}F_{lp}(\upsilon,\varpi) \quad . \tag{12}$$

This equation means translation (ϕ_0, σ) can be determined from the f_{lp} and g_{lp} using phase correlation method [13][15]. Rotation and scale are computed from this translation.

Now, another image f' is mapped from the image f using computed rotation ϕ_0 and scale μ.

$$f'(x,y) = f(\mu(x\cos\phi_0 + y\sin\phi_0), \\ \mu(-x\sin\phi_0 + y\cos\phi_0)) \quad . \tag{13}$$

Translation (x_0, y_0) between the image g and the image f' is computed using the phase correlation method, again.

$$\frac{G(n,m)F'^*(n,m)}{|G(n,m)F'(n,m)|} = e^{j2\pi(nx_0+my_0)} \quad . \tag{14}$$

Eccentricity and its rotation are denoted as,

$$\begin{cases} \theta_\varepsilon = \tan^{-1}(\varepsilon/L) \\ \phi_\varepsilon = \tan^{-1}(y_0/x_0) \end{cases} , \tag{15}$$

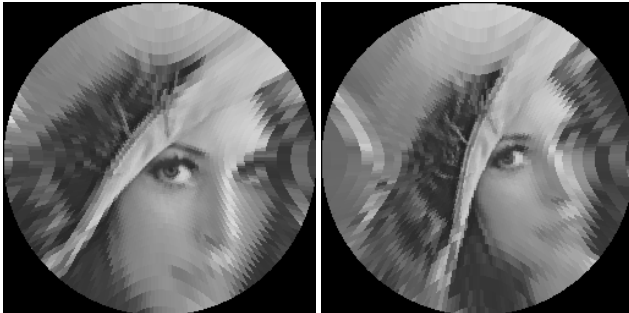where $\varepsilon = \sqrt{x_0^2 + y_0^2}$.

When FMI descriptor is applied for the images f, f, and g, a window function is used. In addition, the following high-pass filter is used for DFTs F and G [13].

$$HPF(n,m) = (1.0 - \cos(\pi n)\cos(\pi m)) \\ * (2.0 - \cos(\pi n)\cos(\pi m)) \quad , \tag{16}$$

where $-0.5 \le n, m \le 0.5$.

(a)uniform-resolution images f and g

(b)undistorted foveated images f_{UDF} and g_{UDF}
Fig.4 An image and a transformed image
using rotation, scale and translation

C. Applying to Space-variant Image UDF

FMI descriptor assumes to be applied for a linear-coordinate (Cartesian) image. However, even if the image *DF* by the fovea sensor is transformed into the linear-coordinate image *UDF* by removing distortion, we need to take non-uniform resolution into account. Multi-resolution images f_j and g_j (where j is resolution level), computed from the images f and g by using Discrete Wavelet Transform (DWT), are shown in Fig.5. Haar wavelet is used for this computation as a mother wavelet. The proposed method applies FMI descriptor for the multi-resolution image from low level to high level, and estimates eccentricity on a reliable level j_m using phase correlation peak as criterion. In addition, the proposed method also defines a variable window function with FOV level λ in order not to sacrifice high accuracy in the central FOV. That is, the higher-level eccentricity is estimated by changing an effective FOV smaller, based on the eccentricity estimated when the window size is the largest. This paper applies Hanning window for the variable window function.

Hanning window:

$$\begin{cases} wnd(x,y) = 0 \quad (\sqrt{x^2+y^2} > L\tan\theta_{\max}\cdot l^\lambda) \\ \quad\quad = 0.5 - 0.5\cos(2\pi\sqrt{x^2+y^2})) \quad , \\ \quad (0 \le \sqrt{x^2+y^2} \le L\tan\theta_{\max}\cdot l^\lambda) \end{cases} \tag{17}$$

Figure 6 shows the image f after applying the window function with different FOV level λ. Figure 7 is a block diagram of the proposed method.

Fig.5 Multi-resolution images f_j, g_j by Haar wavelet

(a)λ=0   (b) λ=1   (c) λ=2
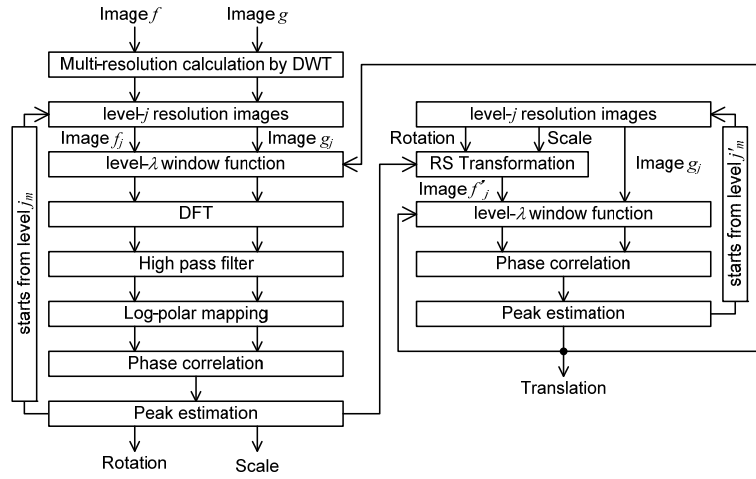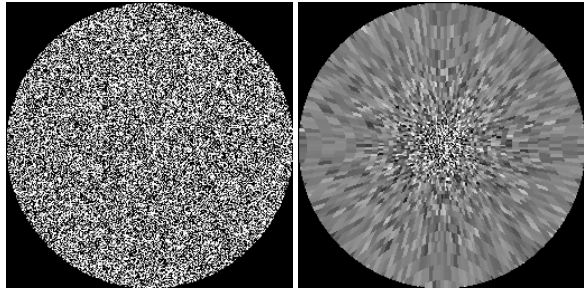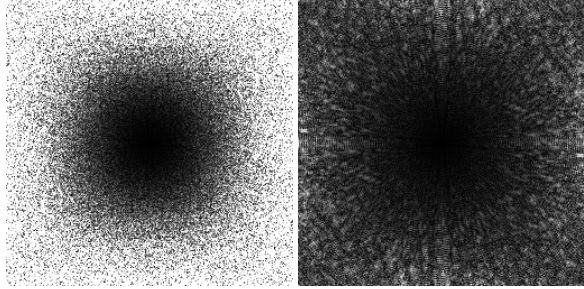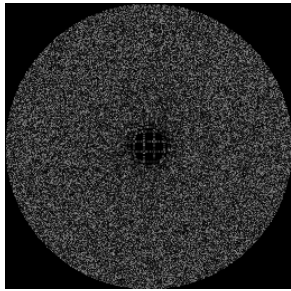Fig. 6 Image f when using different size of a window function

Fig. 7 Block diagram of the proposed method

(a)a white Gaussian noise (WGN) image *w* and its *UDF*-image w_{UDF}

(b)DFTs *W* and W_{UDF} , after high-pass filtering

(c)Foveation noise image w_{fn}
Fig.8 Comparison of a WGN image *w* and its *UDF* image w_{UDF}

III. EXAMINATION

A. *Foveation Noise*

The proposed method uses high frequency components emphasized by high-pass filter for the phase correlation method. However, it is obvious that many of high frequency components in the image of Fig.4(a) are lost in the corresponding image *UDF* of Fig.4(b). This loss causes a risk of wrong estimation by the phase correlation. Thus, in order to reduce this risk, the proposed method estimates the eccentricity by applying FMI descriptor for the multi-resolution image from a lower-level image where resolution is less space-variant.

Figure 8(a) compares a white Gaussian noise image *w* and its *UDF* image w_{UDF}. Figure 8(b) shows that magnitude of their DFTs *W* and W_{UDF} (where the high-pass filter is applied) is influenced by foveation mainly in high frequencies. Figure 8(c) is a foveation noise image w_{fn} that shows absolute value of difference between Invert DFTs of *W* and W_{UDF}. Root Mean Square (*RMS*) of image intensity is used for estimating the foveation noise.

$$RMS = \sqrt{\frac{\left\| w_{fn} \right\|^2}{N_x \cdot N_y}} \ , \tag{18}$$

where $N_x \times N_y$ is an image size of the image w_{fn}.

Figure 9 shows the resolution level versus average of *RMS* of the foveation noise image computed from 100 white Gaussian noise images in each level. This simulation uses AdWAF model in conditions of θ_0= 9.826[°], θ_1=19.107[°], θ_2= 34.715[°], and θ_{max}= 60.0[°] for foveation. In addition, 3 cases of the variable window function with a different FOV level (where a basis *l* is 0.7) are compared with a case without the window function. Consistency between the proposed method and properties of wide-angle fovea sensor (i.e., wide-angle FOV and high resolution in the central FOV) are examined as follows.

1) *RMS* of the foveation noise is zero when the resolution level *j* is less than 4, and increases as the level *j* gets larger. This indicates that the resolution level should be decreased as the eccentricity is estimated in a wider-angle FOV.

2) As the FOV level λ of the variable window function gets larger (i.e., the FOV gets narrower), *RMS* gets smaller. This indicates that the FOV level should be increased as the eccentricity is estimated in the higher resolution.
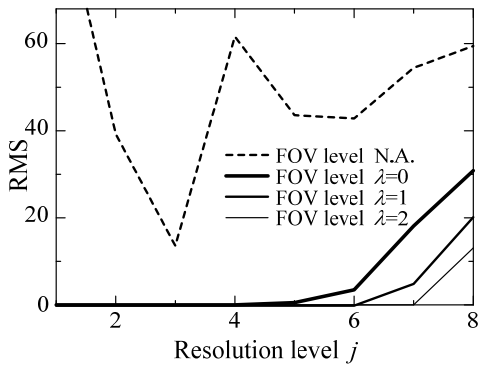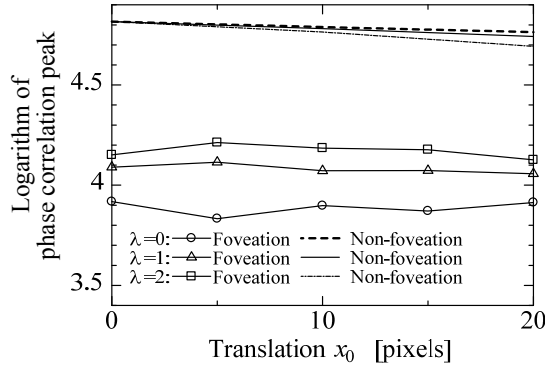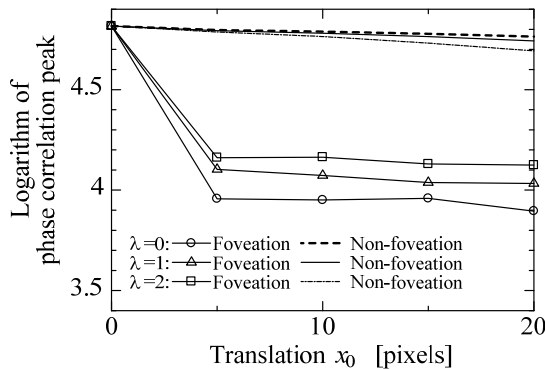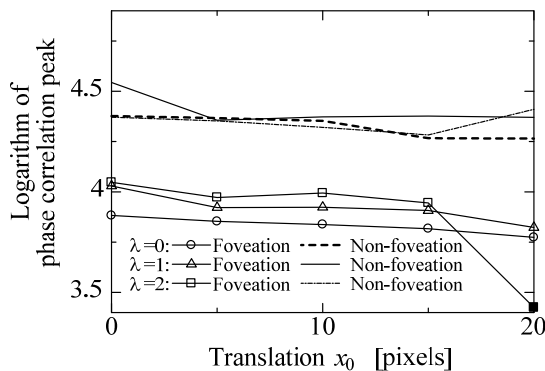
1381

Fig.9 RMS of foveation noise image

(a) when μ=0.8

(b) when μ=1.0

(c) when μ=1.2

Fig.10 Experimental results

B. Experiments

Figure 10 shows experimental results. Being given the translation x_0, logarithms of phase correlation peak determined automatically by the proposed estimation method are compared between the uniform-resolution Lenna image

and its *UDF* image, in 3 cases of the scale μ=0.8, 1.0, and 1.2 (where y_0=0 and ϕ_0=0[°]). In these results, all cases except when μ=1.2 and x_0=20 (shown as a black square) have succeeded in estimating the translation x_0 by an error less than 1 [pixel] (Also, the scale and rotation are estimated accurately). The wrong estimation case mentioned above occurs when the highest FOV level λ=2 (i.e., the narrowest FOV). A local maximal value seems to be selected wrongly because the foveation noise reduces the determined phase correlation peak largely. We note that the higher FOV level gives us the higher value of the phase correlation peak. This means that the high resolution in the central FOV can be preserved by the variable window function, although the high FOV level has a weakness of the translation.

REFERENCES

[1] Schwartz, E.L., Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception, Biological Cybernetics, Vol.29 (1977), pp.181-194.

[2] Sandini, G. and Tagliasco, V., An anthropomorphic retina-like structure for scene analysis, Computer Graphics and Image Processing, 14 (1980), pp.365-372.

[3] Berton, F., Sandini, G. and Metta, G., Anthropomorphic Visual Sensors, Encyclopedia of Sensors, Edited by C.A. Grimes, E.C.Dickey and M.V.Pishko, Vol. 10 (2005), pp. 1-16.

[4] Van der Spiegel, J., Kreider, G., Claeys, C., Debusschere, I., Sandini, G., Dario, P. and et al., A foveated retina-like sensor using CCD technology, Analog VLSI Implementations of Neural Networks. Kluwer, C.Mead and M.Ismail, Boston (1989).

[5] Wodnicki, R., W.Roberts, G. and Levine, M.D., A foveated image sensor in standard CMOS technology, In Custom Integrated Circuits Conference, Santa Clara, California (1995).

[6] Shimizu, S. et.al., Vision Sensor with Wide Angle and High Distortion lens, Video proceedings of IEEE International Conference on Robotics and Automation (1995), Visual Sensing 3.

[7] Kuniyoshi, Y., Kita, N., Sugimoto, K., et.al., A Foveated Wide Angle Lens for Active Vision, Proc. of IEEE International Conference on Robotics and Automation (1995), pp.2982-2988.

[8] Shimizu, S., Wide-Angle Foveation for All-Purpose Use, IEEE/ASME transaction on Mechatronics (published).

[9] Shimizu, S. and Burdick, J., Eccentricity Compensator for Log-Polar Sensor, 2007 IEEE International Conference on Robotics and Automation (2007) , pp.4214-4219

[10] Shimizu, S., Multi-Functional Application of Wide Angle Foveated Vision Sensor in Mobile Robot Navigation, Journal of Robotics and Mechatronics, vol.14, No.4 (2002), pp.382-389.

[11] Shimizu, S., Wide-Angle Vision Sensor with High-Distortion Lens (Detection of Camera Location and Gaze Direction Based on the Two-Parallel-Line Algorithm), JSME International Journal, Series C, Vol.41, No.4 (1998), pp.893-900.

[12] Casasent, D. and Psaltis, D., Position, Rotation and Scale-invariant Optical Correlation, Applied Optics, Vol. 15 (1976.), pp.1795-1799.

[13] Reddy, B.S. and Chatterji, B.N., An FFT-based Technique for Translation, Rotation and Scale-Invariant Image Registration, IEEE Transactions on Image Processing, Vol.5, No.8 (1996), pp.1266-1271.

[14] Chen, Q., Defrise, M. and Deconinck, F., Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 16, No.12 (1994), pp. 1156-1168.

[15] Horner, J. and Gianino, D., Phase-only matched filtering, Applied Optics, Vol.23, No.6 (1984), pp. 812-816.