

A Neural Network Model of Face Detection for Active Vision Implementation

Yasuomi D. Sato^{*,**}, Yasutaka Kuriya^{*}

^{*} Department of Brain Science and Engineering, Graduate School for Life Science and Systems Engineering, Kyushu Institute of Technology, Japan

^{**} Frankfurt Institute for Advanced Studies (FIAS), Goethe University Frankfurt, Germany

Abstract : We suggest an active vision design in a neural style to achieve face detection, or face recognition in gaze, understanding network structures on impaired face recognition in developmental disorders such as autism spectrum deficits and Asperger's syndrome. The core of the active vision design is so-called corollary discharge (CD) on an ascending pathway of the superior colliculus to the frontal eye field (FEF) via the mediodorsal thalamus. The CD is the signal about upcoming eye movements for achieving the stable visual perception. Assuming that the CD is parameterized as controlling dynamical projection creation between the FEF and the Fusiform gyrus, face detection performance is shown to gradually decrease when the value of the CD parameter increases. This indicates a deficit of accurate eye motion, which may elicits impaired face recognition in gaze.

Keywords – Active vision design, Face detection, Corollary discharge, Dynamic link architecture, Elastic graph matching

I. INTRODUCTION

The deficits of face recognition (or face perception) are frequently diagnosed in developmental disorders such as autism spectral disorders (ASDs) and Asperger's syndrome symptoms [1]. They are seen even in schizophrenia patients [2]. The face recognition deficiencies evoke serious social problems on developmental process in the children, because face recognition is one of very significant requirements to acquire social communication skills in childhoods. The urgent task is thus to correctly understand and then solve a mechanism on impaired face recognition. This will be allowed to remove social problems occurred in the patients and enable them to healthy come back to common life.

Fusiform Gyrus (FFG) is regarded as a key brain area in face recognition. A deficit in cholinergic innervations of the FFG was observed in adults with ASD. The deficit may be related to not only current but also childhood impairment of social functioning [3]. But why the cholinergic activity is reduced is still unclear, although the cholinergic activity has been well-known to regulate the function of the visual pathway, including the FFG. To answer this question, we will have to overview a whole configuration of visual pathway, including a visuomotor system and so forth. In addition, we will have to study cognitive modeling of visual perception,

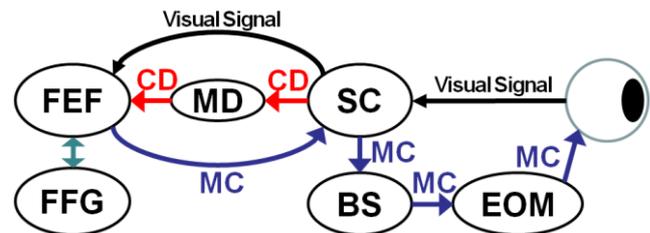


Fig. 1 Two pathways of corollary discharge (CD) and motor command (MC) after visual signals are transmitted to the superior colliculus (SC). The ascending pathway of the CD: The SC to the Frontal eye field (FEF) via the mediodorsal thalamus (MD), which projects to the Fusiform gyrus (FFG). The descending pathway of the MC: The SC to the extraocular muscles (EOM) via the Brain Stem (BS). This figure is referred to [4-6].

In general, an object always and complicatedly moves. The bulks of eye also move such as saccade. Nevertheless, visual perception for moving objects keeps stably reflected in the brain, possibly, in the frontal eye field (FEF)). Such visual stabilities are not accomplished with impaired corollary discharge (CD). In Fig. 1, the CD pathway of the superior colliculus (SC) to the FEF through the mediodorsal thalamus (MD) was empirically confirmed to convey signals about forthcoming eye movements (that is, when a saccade will happen and where it will go) [4]. Such an ascending signal of the CD allows the FEF to feedback motor command signals (called MC in Fig.1) to the SC, to generate a saccade for precise object tracking.

However [5] shows that inactivation in the MD can impair the CD, which makes it very difficult to keep track of an object. Therefore, this physiological experiment indicates that the CD plays an important role in achieving precise and accurate eye track of an object as well.

This is the same of achieving face recognition, keeping track of the face. If the CD can contribute to face recognition in gaze, it leads to understanding deficits of the face tracking (or face attention) by impaired CD pathway. In an fMRI experiment, an intensive connectivity between the FEF and FFG is already observed [6]. If a mechanism on visual face recognition on a network of the FEF and FFG with the CD pathway is clear, it can be expected as helpful to understand mechanisms on face recognition deficits in developmental disorders. For this purpose, it is very important to study a neural model of recognition/detection of a face, including MC signals on CD pathway.

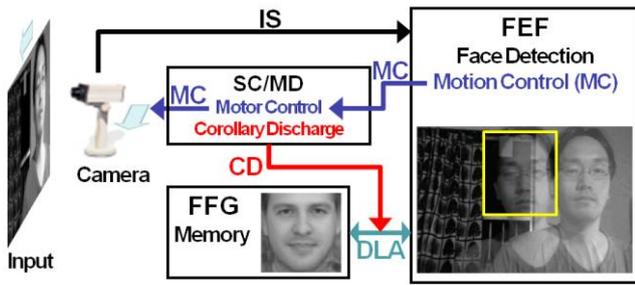


Fig. 2. An active vision configuration for face recognition. Image signals (ISs) are delivered to the FEF from a camera, in which a motion of the face is calculated when detecting a position of the face on an input image. Information about the motion as a motor command (MC) signal descends to the camera via the SC/MD module. It thus allows the camera to smoothly track a face. In the SC/MD module, a parameter of the corollary discharge (CD) controls matching process for face recognition/detection in the dynamic link architecture (DLA). Here shorten forms of the SC/MD, FEF and FFG is referred as in Fig.1.

We propose an active vision design in a neural style for achieving face detection. The dynamic link architecture (DLA) [7-9] employed in this work can frequently be regarded as a powerful tool for explaining two following points: (1) How neural activities in the FEF can be projected to the ones in the FFG, including the ascending pathway of the CD. (2) How face recognition (or detection) can be achieved. Assuming that the CD plays an important role in dynamical creation of projections of the FEF to FFG, we show face detection performance increased with changes of the CD as a control parameter. Within a framework of the active vision design, we will discuss considerable mechanisms on face recognition deficits observed in patients with developmental disorders.

II. ACTIVE VISION DESIGN

2.1. System Configuration

Let us begin by looking at a whole system for achieving face recognition in gaze. This mainly consists of three

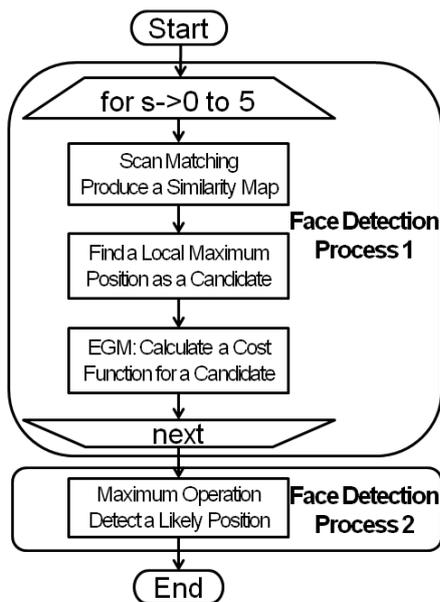


Fig. 3. A flowchart of face detection process. It consists of two sub-processes. The first (Face Detection Process 1) is candidate position finding while the second (Face Detection Process 2) is the most likely position detection.

modules, for the simplicity, the FFG, the FEF, as well as the SC/MD (see Fig.2). The FFG as memory stores a facial image (called the M). In the FEF, a face is detected on an input (I) image, which is sent as an image signal (IS) by a camera. The face detection is done by matching the images I to M within a framework of the DLA. The matching process is controlled by the CD parameter in the SC/MD. Motion is estimated in comparison with the previous frame. It is delivered as a motor command (MC) signal to a servomotor in the camera through the SC/MD. Thus, the camera can track a face.

2.2. Feature Representation in FFG module

We employ an image of the average face created by a face generator in [10], whose size is $A_0=60 \times 60$ pixels. Since feature representations are composed of the different scales, the image is down-sampled with a_0^s ($s=0, \dots, 5$), called the M_s . Here $a_0=0.85$ and $s=0$ implies the original size of the image M . A square graph of $(n \times n - 4)$ nodes without any vertexes is set on each resolution image M_s . Here n ($=5$) is the number of full nodes on a row and column of the square graph. Each node on the image M_s is convoluted with a family of Gabor functions $\Psi_r(z)$. r ($=0, \dots, 7$) is an orientation parameter. The Gabor feature usually consists of 8 different orientation components. Each orientation component is given by a convoluted value ($\hat{J}_r^{M_s}$):

$$\hat{J}_r^{M_s}(p_s) = \int F^{M_s}(p_s - \hat{p}_s) \Psi_r(p_s - \hat{p}_s) d^2 \hat{p}_s, \quad (1)$$

$$\Psi_r(p_s) = \frac{k_r^2}{\sigma^2} \exp\left(-\frac{k_r^2 p_s^2}{2\sigma^2}\right) \left[\exp(ik_r p_s) - \exp\left(-\frac{\sigma^2}{2}\right) \right]. \quad (2)$$

$\sigma=2\pi$. The wave number vector k_r can be expressed as

$$\vec{k}_r = \begin{pmatrix} k_{rx} \\ k_{ry} \end{pmatrix} = \begin{pmatrix} k \cos \varphi_r \\ k \sin \varphi_r \end{pmatrix}, \quad k = \frac{\pi}{2}, \quad \varphi = \frac{\pi}{8} r. \quad (3)$$

One orientation component $J_r^{M_s}$ in the Gabor feature takes an absolute value of $\hat{J}_r^{M_s}$:

$$J_r^{M_s}(p_s) = |\hat{J}_r^{M_s}(p_s)|. \quad (4)$$

We mention physiological backgrounds of multi-scale feature representation for a face. Physiological experiments for receptive field already finds that neurons being tuned to high spatial frequencies have narrower tuning range than neurons being tuned to low spatial frequencies. Also, the receptive field structure represented by Eq. (2) is observed as constructed by multiplying a global sinusoidal grating by a bell-shaped Gaussian envelope [11].

Furthermore, human vision can be conceived to be achieved through low pass filter processing [12]. Once an input image about some environmental scene is received on a retina, the highest spatial frequency component of the Gabor filter is sequentially discarded in bottom-up flow (referred to [13]). The prospective discarding spatial frequency elements may be stored in another area through another pass in the visual cortex. The FFG can be considered as one of candidates. Therefore, the multi-scale Gabor feature representations are suitable for modeling of memory.

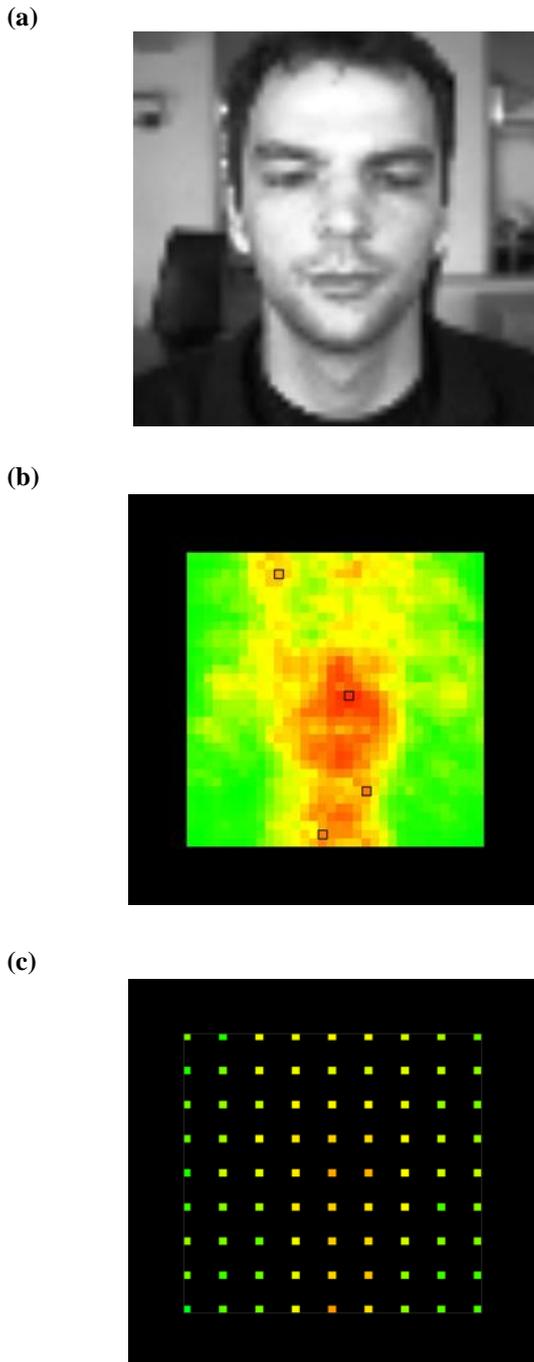


Fig. 4. Similarity maps when a static grayscale image of a face, used as an input image (a), is scanned with a square graph at M_s . (b) $n^{cd}=1$ and (c) $n^{cd}=4$ where n^{cd} is the scanning step number. In (b) and (c), black squares show a local maximum position on the similarity map.

2.3. CD controlled Graph Matching in FEF module

In analogy, each frame on the input (I) image is convoluted with a family of Gabor function. The main face detection processing is then proceeded. As shown in Fig.3, the face detection processing consists of two sub-processes of candidate position finding and the most likely position detection.

2.3.1. Scan Matching

In the first sub-process, in order to pick up candidate positions, an entire or fragmentary similarity map for each scale M_s is calculated with scan matching of the undistorted graph G_s onto the image I :

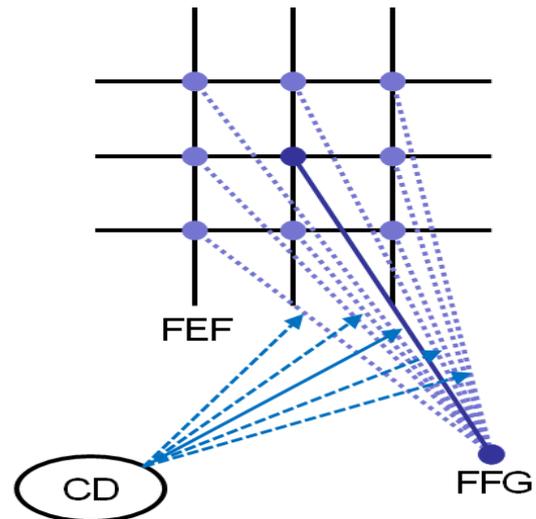


Fig. 5. Dynamic link architecture (DLA) between the FEF and FFG. One feature detector for model representation on the FFG, which has been projected to all of feature points on the FEF, then finds the optimal point that takes the highest similarity to the model representation through the CD control (see a solid line and a solid arrow). In the CD, the other controllers (broken arrows) weaken projections between the FEF and the FFG (broken lines).

$$e_s^S(x_I) = \frac{1}{|G_s|} \sum_{p_s \in G_s} \frac{\sum_r J_r^I(x_I) \cdot J_r^{M_s}(p_s)}{\sqrt{\sum_r (J_r^I(x_I))^2 \sum_r (J_r^{M_s}(p_s))^2}}, \quad (5)$$

where $e_s^S(x_I)$ is the similarity of the Gabor feature at x_I on the image I to the one for each node p_s . It is noticed that G_s sometimes represents a set of nodes of the graph. In the scan matching process, the left-upper of the image M_s on which a square graph is being set up, is firstly adjusted to the left-upper of the I . The M_s is repeatedly scanned to the right for each row on the I . The scan ends when the right-lower of the M_s arrives at the right-lower of the I (not shown here). As a result, we obtain a similarity map at each image M_s . We exemplify similarity map results when one facial image I is scanned with a square graph at the level M_s (Fig. 4).

Next, we pick up some candidate positions that can be expected as the one of a face. Here let the candidate position x_I^c be defined as the center of the image M_s or of its square graph. It corresponds to the local maximum on the similarity map. The local maximum is satisfied with a gradient condition when all differences of the x_I^c to its nearest neighbors x_I^n take positive:

$$e_s^S(x_I^c) - e_s^S(x_I^n) > 0. \quad (6)$$

The candidate pixel is depicted with a black square on the similarity map as shown in Fig. 4(b).

2.3.2. Elastic Graph Matching (EGM)

The EGM for each candidate position x_I^c is computed to obtain the maximum value of the cost function E_s^c , which is given by

$$E_s^c(x_I) = e_s^S(x_I) - \lambda_d e_d^d(x_I), \quad (7)$$



Fig.6. Facial images of a same or different person with the background in a Bio ID dataset of 1521 images.

$$e_s^d(x_I) = \frac{1}{|G_s|} \sum_{p_s \in G_s} \frac{\left| \sum_{p'_s \in G'_s} D_{p_s p'_s}^{M_s} - \sum_{p'_s \in G'_s} D_{x_I p'_s}^I \right|}{\left| \sum_{p'_s \in G'_s} D_{p_s p'_s}^{M_s} \right|}, \quad (8)$$

$$+ 1 - \frac{1}{|G_s|} \sum_{p_s \in G_s} \frac{\left| \sum_{p'_s \in G'_s} A_{p_s p'_s}^{M_s} \cdot \sum_{p'_s \in G'_s} A_{x_I p'_s}^I \right|}{\left| \sum_{p'_s \in G'_s} A_{p_s p'_s}^{M_s} \right| \cdot \left| \sum_{p'_s \in G'_s} A_{x_I p'_s}^I \right|}$$

where $e_s^d(x_I)$ represents the elasticity of the graph on the image I . λ_d is a constant parameter for the graph elasticity. $\lambda_d=0.05$, except for obtaining a similarity map $E_s(x_I)$ when $\lambda_d=0$. G'_s is a set of nearest neighbor nodes p'_s for p_s . $D_{p_s p'_s}^{M_s}$ and $D_{x_I p'_s}^I$ are the Euclidean distance between nodes p_s (or x_I) and p'_s on the graph of the image M_s (or I). $A_{p_s p'_s}^{M_s}$ and $A_{x_I p'_s}^I$ take one vector form consisting of 4 elements. Each element is an angular between two nearest neighbors on each quadrant, centered at p_s .

Each node on the image I , which corresponds to the node p_s^c of the square graph M_s , surveys an optimal pixel p_s^I taking a maximum of the cost function E_s^c within a search region R :

$$p_s^I = \max_{x_I \in R} \{E_s^c(x_I)\} \quad (9)$$

where R is a set of pixels that can be picked up in a square with the size of $(2q+1) \times (2q+1)$, centered at the pixel corresponding to p_s^c . $q=4$.

The optimal pixel x_m^I is singled out with the maximum operation of all candidates, which must be the most invariant to feature representation for an M face:

$$x_m^I = \max_{c \in C} \{E_s^c(p_s^I)\} \quad (10)$$

Here C is a set of candidates.

2.4. Motor Command

Finally, let us explain briefly an algorithm of face tracking by controlling a camera. In this algorithm, motions of a face are defined as a difference of the face position on the $I^{(i)}$ to



Fig.7. Face detection demonstration when we used a camera. Numerals on the left-upper are respectively the maximum value of a cost function and computational cost value for each frame.

the one on $I^{(i+1)}$

$$x_m^{I^{(i+1)}} - x_m^{I^{(i)}} \quad (11)$$

where i th is the frame index. The information about motions of a face is transmitted to a motor controller in the servo camera via the SC/MD module. The camera can be expected to track a face on the input.

A goal of this article is to propose active vision architecture to understand a mechanism on impaired face detection. It is naturally expected that the accuracy of motion gets worse when face detection performance decreases. Therefore, we omit the concrete explanation about the motor control algorithm in this article.

2.5. Modeling of Corollary Discharge

A role of the corollary discharge (CD) pathway is reconsidered for interpreting the face detection system in a neural style. As mentioned in the Introduction, deficits of the CD decrease the accuracy of eye movements. The CD is thus necessary to control precisely motions of the eye. To increase the accuracy of motion, the accuracy of the face detection also increases.

For this, the core of face detection with scan matching and EGM is feature-based correspondence finding within a framework of the DLA. As shown in Fig.5, one feature detector on the M graph, projected to all pixels in the I , tries to find the optimal with topographic mappings. It also enables us to find the highest invariance to feature representations for a face. This is the essential of searching the local maximum of a similarity map. Also, the same is finding the maximum of a cost function.

However there is still a problem in searching the local maximum, which detection fault is increased if a number of the local maximum is found as shown in Fig. 4(b). As one possibility to solve this problem, we do not create a whole similarity map, but the local similarity map as shown in Fig. 4(c). This can be expected to improve the face detection fault.

Return to neural model interpretations of our face detection system. The scanning step number n^{cd} controls projection patterns. The all-to-all projection pattern is, for

TABLE I
FACE DETECTION PERFORMANCE

Scan Step	The number of correct detected images	The detection rate (%)
$n^{cd}=1$	1485	97.6
$n^{cd}=2$	1474	96.9
$n^{cd}=3$	1472	96.7
$n^{cd}=4$	1445	95.0

A face detection ability test using the database of the BioID with 1521 facial images. An undistorted graph is scanned with scan step n^{cd} on the input image.

example, changed to the sparse when the n^{cd} increases. It addresses that the CD controls neural connectivity patterns between the FEF and FFG. The CD control parameter, namely, the scanning step number is one of the remarkable points in our detection system. The face detection rate computed with our detection system is related to the accuracy of the face detection. This will thoroughly be discussed in the next section.

III. DETECTION PERFORMANCE

We test an ability of our face detection system, using the Bio ID database [14] that involves 1521 facial images. Fig. 6 shows some of facial images in the Bio ID database. On the other hand, an image of an average face produced with many German facial photos in [10] is employed as the model (M) image. This indicates that our face detection system has recognition ability for a face as the object, not personal identity.

In the detection ability test, we do 4 trials of $n^{cd}=1, 2, 3$ and 4. For each trial, the M image should be tried to match to each facial image in the Bio ID database. Correct face detection can be defined when eyes are in 3 of 5th from the top of detected square area. The correct face detection rate is the accuracy of the face detection in this work. The correct detection results are shown in TABLE I.

In TABLE I, when the n^{cd} is increased 1 to 4, the face detection rate is gradually decreased 97.6% to 95.0%. When the scan step number is furthermore increased, the face detection rate becomes decreased (as not shown here), because the number of candidates is decreased.

These results indicate as follows: The first can expect more improvement of face tracking in active vision. In fact, we use a camera. Our face detector can, in real time, capture a face without any severe problems even though the size of the face is bigger or smaller (see Fig. 7). The second is the validity of neural model interpretation of the corollary discharge. Increases of the scan step represent because our model supports experimental results for a functional role of the corollary discharge.

IV. DISCUSSION

We propose active vision architecture in a neural style, assuming that this architecture contains a mechanism to control a motor in a servo camera. The key is the bottom-up signal of the MD to the FEF for information about forthcoming eye movement, which is called the CD, on which the top-down signal is transmitted as a motor

command. The functional role of the CD is to improve or preserve eye movement accuracy.

In this article, making full use of the concept of the DLA, we parameterize the role of the CD for its simplicity, and also show that face detection performance calculated here, which qualitatively supports to an experimental result, is decreased with an increase of the CD parameter value. It is not doubtful that the CD is a necessarily function to object tracking or visual attention.

It is still unclear if or not the impaired CD pathway is a central mechanism on deficits of face recognition. This is due to a decline of face detection performance indirectly means difficulties of face tracking by motor controls. Sect. I Introduction already reports that patients with deficits of face recognition can difficultly pay attention to, or track a face. It can thus be inferred that they are not in healthy communication environment in their childhood so that the CD pathway is impaired in the brain. This may cause a decline of abilities for attention to and track a face to induce deficits of face recognition.

As shown in Fig. 5(c), we have created a sparsely distributed map of the similarity. Since it means a decrease of correlations or connectivity intensities between neural activities of the FEF and FFG, the localized similarity map result may support inactivation of the cholinergic system in the FFG. However, to confirm support of the computational model to the cholinergic inactivation, we will have to study more declines of neural activities in FEF and MD, together with cholinergic inactivation in the FFG in the experiment. Such experiments are not yet reported as far as we know.

There are still lots of improvements in our face detection algorithm. The first priority is the CD deficit parameter n^{cd} . One of goals in this work is to understand functional mechanisms on achievement of face recognition in developmental process. We have to study neural network modeling of the related learning process. For this, physiological experiments that have been done by Bell et al. [15] are of great use as a reference.

Bell et al. addressed that representations of recent sensory input that followed motor commands are stored and updated through anti-Hebbian plasticity at synapses between corollary discharge conveying fibers and the sensory area. The updating is prevented by no motor commands and no plastic changes. If such a learning model is achieved for being implemented into our face detection system with the EGM, this can expect to lead to solving mechanisms on how face recognition can be impaired. Also, it gives us one of solutions that the deficit of the CD harms not only to the FEF activation, but also to updating and storing input image in the FEF. Correspondingly the deficit elicits the cholinergic inactivation in the FEF. Furthermore, the motor command signals can no longer be transited for accurate eye movements so that eye tracking or attention to a face is impossible. Therefore, modeling of the learning on the CD pathway is crucial for giving us explanations of symptoms in face recognition disorders.

Taking a glance at a current research field of developmental robotics, researches and developments on physiologically inspired active vision and its relevant learning algorithm does not yet seem to be reported [16]. The learning model implemented into cognitive

developmental robots [17] is the already existing. It cannot be in a neural style, much less modeling of a corollary discharge mechanism. In [18], a multiple forward model for a corollary discharge mechanism is used to computationally simulate experiments on attribution of own actions to intention of self or others. Such a corollary discharge model is still conceptual or cognitive, is not based on the physiological principles, is already done on computational simulations, and is not yet applied into robots.

There is ample scope for some progresses of computer vision technology in a face detection system we proposed as a neural style. Even in the improvement [19], our face detection system has higher face detection performance and comparably rapid computational speeds, compared to the Viola-Jones face detection algorithm [20]. The details are removed here, because they are different directions to this work. But, in the near future, real-time visual information processing models will be achieved on computers. They can be competed with real visual processing in the brain.

V. CONCLUSION

In this article, we studied a neural network model for correct face detection, to understand functional mechanisms on impaired face recognition in developmental disorders such as autism spectrum deficits and Asperger's syndrome. The core of the design is a so-called corollary discharge (CD) mechanism on an ascending pathway of the superior colliculus to the frontal eye field (FEF) via the mediodorsal thalamus. The CD is the signal about upcoming eye movements for achieving the stable visual perception. We assume that the CD plays an important role in dynamical creation of projection of the FEF to the Fusiform gyrus and then show that face detection performance decreases with changes of the CD parameter's value. This can predict a decline of the accuracy of eye motion. From a result of the detection performance decline, considerable mechanisms on the face recognition deficit were discussed, in terms of difficulties of face tracking. We indicated that an active vision design proposed here was neurally plausible as well as powerful to explain deficiencies of face recognition.

REFERENCES

- [1] M. L. Phillips, Facial processing deficits and social dysfunction: how are they related? *Brain*, 127(8), 2004, 1691-1692.
- [2] F. Irani, S. M. Platek, I. S. Panyavin, M. E. Calkins, C. Kohler, S. J. Siegel, M. Schachter, R. E. Gur, R. C. Gur, Self-face recognition and theory of mind in patients with schizophrenia and first-degree relatives. *Schizophrenia Research* 88, 2006, 151-160.
- [3] K. Suzuki, G. Sugihara, Y. Ouchi, K. Nakamura, M. Tsujii, M. Futatsubashi, Y. Iwata, K. J. Tsuchiya, K. Matsumoto, K. Takebayashi, T. Wakuda, Y. Yoshihara, S. Suda, M. Kikuchi, N. Takei, T. Sugiyama, T. Irie, Reduced Acetylcholinesterase Activity in the Fusiform Gyrus in Adults With Autism Spectrum Disorders. *Arch. Gen. Psychiatry* 68(3), 2011, 306-313.
- [4] M. A. Sommer, R. H. Wurtz, What the Brain Stem Tells the Frontal Cortex. I. Oculomotor Signals Sent From Superior Colliculus to Frontal Eye Field Via Mediodorsal Thalamus. *J. Neurophysiol.* 91, 2004, 1381-1402.
- [5] M. A. Sommer, R. H. Wurtz, What the Brain Stem Tells the Frontal Cortex. II. Role of the SC-MD-FEF Pathway in Corollary Discharge. *J. Neurophysiol.* 91, 2004, 1403-1423.
- [6] J. Xaing, J. Chen, The study on Functional Connectivity between Frontal Eye Field and Other Brain Cortex in Visual Processing. *2010 Int. Conf. on Computational Aspects of Social Networks (CASoN)*, 2010 333-336.
- [7] C. von der Malsburg, The correlation theory of brain function. *Internal Report, Max-Planck-Institut für Biophysikalische Chemie*, Postfach 2841, D3400 Gottingen, FRG, 1981.
- [8] C. von der Malsburg, How are nervous structures organized? in Synergetics of the Brain, *Proc. Int. Symp. Synergetics*, E. Basar, H. Flohr, H. Haken, and A. Mandell, Eds. Berlin, Germany: Springer, 1983, 238-249.
- [9] C. von der Malsburg, Nervous structures with dynamical links. *Berichte der Bunsengesellschaft für Physikalische Chemie* 89, 1985, 703-710.
- [10] <http://www.faceoftomorrow.com/>
- [11] M. Čadik, Human Perception and Computer Graphics. Czech Technical University Postgraduate Study Report, 2004.
- [12] N. Thibos, Image processing by the human eye. In: Visual Communications and Image Processing IV, William A. Pearlman (ed), *Proc. SPIE 1199*, 1989, 1148-1153.
- [13] Y. D. Sato, Y. Kuriya, Visual Constructed Representation for Object Recognition and Detection. In B.-L. Lu, L. Zhang, and J. Kwok (Eds.): *ICONIP 2011, Part III, LNCS 7094*, 611-620, Springer-Verlag Berlin Heidelberg, 2011.
- [14] O. Jesorsky, K. Kirchberg, R. Frischholz: Robust Face Detection Using the Haus-dorff Distance. In J. Bigun and F. Smeraldi editors, *Audio and Video based Person Authentication - AVBPA 2001*, Springer, 2001, 90-95.
- [15] C. C. Bell, A. Caputi, K. Grant, J. Serrier, Storage of a sensory pattern by anti-Hebbian synaptic plasticity in an electric fish. *Proc. Natl. Acad. Sci. USA* 90, 1993, 4650-4654.
- [16] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, C. Yoshida, Cognitive Developmental Robotics: A Survey. *IEEE Transactions on Autonomous Mental Developmental* 1(1), 2009, 12-34.
- [17] A. Watanabe, M. Ogino, M. Asada, Mapping facial expression to internal states based on intuitive parenting. *Journal of Robotics and Mechatronics* 19(3), 2007, 315-323.
- [18] M. Otake, K. Arai, M. Kato, T. Maeda, Y. Ikemoto, K. Kawabata, T. Takagi, H. Asama, Experimental Analysis of the Attribution of Own Actions to Intention of Self or Others by the Multiple Forward Models. *Journal of Robotics and Mechatronics*, 19(4), 2007, 482-488.
- [19] Y. D. Sato, Y. Kuriya, Efficient and Effective Gabor Feature Representation for Face Detection. *International Journal of Computer and Information Engineering* 6, 2012, 9-12.
- [20] P. Viola, M. J. Jones, Robust Real-Time Face Detection. *Int. J. Computer Vision* 57(2), 137-154, 2004.