Multi-Aspect Approach to Robotic Anticipation for Supporting Human’s Multiple-Intelligence in Natural Scene

Kohji Kamejima
Faculty of Information Science and Technology, Osaka Institute of Technology
1-79-1 Kitayama, Hirakata 573-0196 JAPAN
e-mail: kamejima@is.oit.ac.jp

March 4, 2010

Abstract: A multi-aspect approach is applied to one-step anticipation for tightening human-robot interaction in naturally complex scene. Via symbolic massage exchange, human’s linguistic intelligence identifies possible solutions by robotic intelligence. In reference to signal space as operable representation of the scene, human’s body-kinesthetic intelligence and robotic control systems are co-oriented. As a common basis of these interaction aspects, in this paper, visible class of ‘anticipants’ is introduced on multi-viewpoint imagery. Despite not-yet-structured representation, the anticipants focalizes robotic intelligence to the scope of human’s subsequent performances.

Keywords: Multi-Aspect Modeling; Human-Robot Interaction; Robotic Anticipation

Table of Contents

1 Introductory Remarks 2
2 Existence of Perceptual Invariance 2
3 A General Framework 4
4 Information Intensive Maneuvering – Randomness Based Scenario 5
5 Computational Aspect of Anticipants 7
6 Combinatorial Aspect of Anticipants 7
7 Control Aspect of Anticipants 8
8 Square Root LV-System 9
9 Discussions 11
10 Concluding Remarks 12

---

1 Introductory Remarks

On vehicles, human’s physical mobility exceeds by far the capability for understanding the situation based only on inherent perception. Noticing that decision support ‘one second’ prior to the critical situation is enough to evade 90% of fatalities, the scope of robot vision should be expanded to interactive support of human’s anticipative capability [37], [27]. By combining ‘digital roadway map’ and machine vision, in fact, the horizon of human’s perception is expanded from visibles to informatic vicinity where possible situations can be computed prior to ‘physical time’ [16]. Despite rapid progress in computer science and network technology, however, continual update of ‘digital roadway map’ to as-is scene is still an open question.

By substituting information generated by today’s robotic systems for human’s decision process, mental load can be reduced in complex situations where awareness of objects easily fills up the capacity of concentration. Intended for effective support, thus, the concept of information intensive maneuvering has been expanded to vehicle-highway integration. In search of information basis of mobile robots, it has been demonstrated that essential capability for maneuvering arises from location specific integration of multi-viewpoint imagery; site-route graph in bird’s-eye view is associated with landmark views to provides logical anticipation of scene [19]; inter-pattern kinetics is introduced in scene image as a computer implementation of subconscious part of perceptual cycle [29]. The central concept of information intensive maneuvering has been effective in autonomous navigation from early mobile robots [38] to latest unmanned vehicles [33].

Recent advancement of information technology makes it possible to expand the basis of decision to in-situ satellite-roadway-vehicle network as shown in Fig. 1. By re-constructing key technologies of robot vision, it has been demonstrated that geometric descriptions of objects are interactively compiled into as-is ‘digital map’ [32]. In this paper, the schematics of the interactive mapping is reformulated within the framework of multi-aspect design [36]. Via the multi-aspect approach, we can integrate the diversity of perception-decision processes without a priori definition of ‘robotic intelligence’.

2 Existence of Perceptual Invariance

From the viewpoint of Gibson [6], we are surrounded by optical flux modulating complete information that can afford to induce spontaneous processes preestablishingly. Such affordance should be captured nondeterministically in perceptual cycle [29] as shown in Fig. 2; the understanding of the world is interactively generated based on the innovation arising in the discrepancy between observation and prediction by resulted schema. Basic idea of stochastic systems theory [11] implies that the innovation provides for ‘unbiased’ decision [24], [21]. The stability of the perceptual cycle is supervenient to something invariant in the real world; 3D geometry of objects is invariant with respect to self-location [3]; linguistic structure provides identical impression to object images [2]; mathematical complexity of optic array is decoded to recognize natural scene [25]. Despite in-depth investigations since the dawn of history, the something-invariant is still left to an irreducible information.

Being embedded in the invisible invariance of the world, human’s perception is equipped with the
capability for situation understanding through the cross-reference of mutually independent aspects. As a revelation of the multi-aspect access, consider maneuvering process along a town street as shown in Fig. 3. From this image, we can extract a simple perceptual rule: proportionality of patterns randomly distributed on the roadway as illustrated in Fig. 4 [13]; by indexing the distance to an object in terms of the height in the image, a kind of perceptant, the size of roadway pattern is linearly reduced according to the increase of the distance. The perceptual linearity in man-readable patterns can be extracted from ‘insensible’ information as well. For example, figure 5 visualizes the transition of the power spectrum associated with ‘line images’: a representation of scale information with respect to the perceptual distance [15]. As shown in this fugue, the linearity is verified via detecting a vanishing point $y_\infty$ at which the spatial frequency of the image diverges. It should be noted that Laplacian-Gaussian channels can filter scale information from brightness distribution in the image plane [26], [10]. Thus, the linear scale shift provides robust aspect of perceptual invariance.

Recent advancements in fractal geometry introduces new aspect of perceptual invariance. Following the collage theorem [1], the expansion of the area supporting the perceptual linearity evokes a mental dynamics as indicated in Fig. 6. By identifying the aggregation of random motion towards the vanishing point with a self-similarity process, the complexity of object images can be coded in terms of simple 2D mappings. In fact, visual perception of animals was shown to exhibit definite sensitivity to computational simplicity arising in the self-similarity rather than iconic simplicity [28].

Furthermore, the ‘whiteness’ of daylight can be stable reference for adjusting trichromatic system
against spectrum shift. In fact, chromatic coherency of natural scene is restored based only on a vestige of daylight contained in illuminant [8].

3 A General Framework

Within the framework of multiple intelligence [5], information for intentional maneuvering should be generated through tight interaction with spatial and body-kinesthetic intelligence as shown in Fig. 7; the physical aspect of maneuvering process is grounded on visibles in real world [7] via logical-mathematical intelligence; by ‘reading’ the physical processes as a space of consistent signals, robotic intelligence generates an aggregation of symbol; the symbol space governed by logical causality [30] induces anticipative visualization of the scene beyond the horizon of physical time.

The establishment of tight linguistic-physical interaction plays a crucial role in various scopes of human-robot cooperation including user’s expectation [31] and active situation understanding [34], e.g. Being supported by neural dynamics, linguistic intelligence \textit{a priori} focalizes the arrival of triggering stimuli and \textit{a posteriori} acknowledges the completion of decision program [22], [23]. This mental expansion implies the existence of latent anticipation scheme under the constraints of ‘physical time’ as shown in Fig. 8; the linguistic understanding process is accessible to entire the stimuli against the time delay.
for neuronal adequacy and reactive to parallel decision sequences towards acceptable decision. For a priori adaptation to consistent results, the linguistic process expands the dimensionality of the ‘present time’ into a space of object symbols [4]. The latent process, whereas, refers to physically clocked signal dynamics so as to evoke the illusion of ‘another possibilities’ [35].

To extend the anticipation scheme from ‘personal world’ to human-robot interaction, in what follows, we introduce a cognitive entity called anticipants on the scene image as illustrated in Fig. 9; on the photocopy of real world, the anticipants accept sketch of route from the multi-viewpoint map; scene image is of sufficient ‘geometric depth’ for accepting ‘agreeable path’ to be generated by machine planning under the control of human’s body-kinesthetic intelligence; via 2D restriction of 3D world, simultaneously, human’s a posteriori intention is coded as the basis of computational recognition process. Thus, robotic recognition-planning process coexists with human’s spontaneous decision towards agreeable scene.

4 Information Intensive Maneuvering – Randomness Based Scenario

Regardless of inherent intelligence or artificial one, maneuvering process maintains tight interaction with complex image of a scene through various aspects as shown in Fig. 10; first, knowing skew caused by perspective projection, vision system expands the boundaries of the roadway from the position of the present. By coding the expansion in terms of fractal attractor [28], next, the ‘one step expansion’ is applied to satellite image for bird’s-eye segmentation of roadway pattern; as the results of the segmentation, finally, the informatic vicinity is indicated on as-is imagery.

In this scenario, the complexity of scene and satellite images are analyzed in terms of Kolmogorov indexes [20]: combinatorial structure on chromatic information, capturing probability in scale space and program length in fractal coding. By using the Kolmogorov complexity, we can introduce ‘imaging process’ itself as the object to be computed in stead of objects to be visualized. Such process level representation of unstructured scene provides formal basis for open logic machine [18] in which the scene is identified in terms of fixed points associated with computational-geometric dynamics.

Suppose that the vehicle is oriented by a vector v on the ground as indicated in Fig. 12. From the
viewpoint of ecological optics [7], a priori basis for spontaneous mobility is provided by random texture mapped onto an image plane $\Omega = X \times Y$. Let $\omega \in \Omega$ be a pixel with brightness distribution $f(\omega)$, and suppose that objects in the scene are visualized as subsets of $\Omega$. Without serious loss of generality, the view of the object to be detected is uniquely represented by a image $\Lambda$ and identified within $\mathcal{F}$: the Borel field generated by the subsets of $\Omega$. In the totality of object images, then, we can introduce basic probability space $(\Omega, \mathcal{F}, P)$ where $P(\omega)$ is uniform probability measure assigned to pixels $\omega$. 

Figure 10: Cooperative Decision in Informatic Vicinity

Figure 11: Open Logic Machine

Figure 12: Intentional Coordinate System [15]
5 Computational Aspect of Anticipants

Let $\mathcal{P}$ be a given class of programs $\pi$ with essential length $\ell(\pi) < \infty$ and consider pixel-wise ‘coloring’ process $\phi(\pi, \theta)$ with control parameter $\theta$ selected in a preassigned set $\Theta$. By identifying generalized imaging mechanism with the ‘random variable’ $\phi(\pi, \theta)$, we have the following Ejiri’s Paradox [12]:

**Proposition 1** Define Kolmogorov complexity of generalized imaging process by

$$K_\phi(\mathcal{F}|\Theta) = \min_{(\phi(p, \theta) \in \mathcal{F})} \ell(p).$$

Then, we can select $\Lambda \in \mathcal{F}$ such that $\phi(\pi, \Theta) \neq \Lambda$ for any $\pi$ and $\theta$. This implies $K_\phi(\mathcal{F}|\Theta) = \infty$; i.e., recognition of general patterns is undecidable.

As a natural restriction, let the target of imaging process be coded by a class of self-similar patterns satisfying

$$\Xi = \bigcup_\nu \mu_i(\Xi), \quad \mu_i(\Xi) \cap \mu_j(\Xi) = \emptyset, \quad (1)$$

where $\nu = \{ \mu_i \}$ denotes a set of contraction mappings from $\Omega$ into itself. By combining convergence theorem [9] and collage theorem within the probability space $(\Omega, \mathcal{F}, P)$, we can design the mapping set $\nu$ to approximate any pattern in $\mathcal{F}$ in the following sense

$$K_\phi(\Xi|\nu) = \log_2 \|\nu\| + \frac{1}{2} \log_2 (2\pi e)^2 \sum_{i} |\Sigma[\mu_i(\Xi)/\Omega]| < \infty,$$

where $\|\nu\|$ and $\Sigma[\cdot]/\Omega$ denote the size and relative variance of a set $\cdot$, respectively [14]. Furthermore, the self-similarity (1) can be visualized as follows;

**Proposition 2** For $\mu_i \in \nu$ to be selected with probability $p_{\mu_i}$, there exists an measure $\lambda^p_\Xi$ on $\mathcal{F}$ invariant with respect to the following Markov operation:

$$\lambda^p_\Xi(\cdot) = \sum_{\mu_i \in \nu} p_{\mu_i} \chi^p_\Xi[\mu_i^{-1}(\cdot)], \quad (\cdot) \in \mathcal{F}.$$ 

For unknown $\nu$, we can evaluate the probability for capturing $\Xi$ to be generated as the solution to the following partial differential equation:

$$\frac{1}{2} \Delta \varphi_p(\omega|\nu) + \rho[\chi^p_\Xi(\nu) - \varphi_p(\omega|\nu)] = 0. \quad (2)$$

Due to the random selection of $\mu_i \in \nu$, the capturing probability $\varphi_p(\omega|\nu)$ is determined within complexity factor $\rho = \log \|\nu\|$; by this freedom for the design of mapping set $\nu$, we can induce the following conditional probability

$$dP(\omega|\nu) \propto \varphi_p(\omega|\nu)dP(\omega). \quad (3)$$


6 Combinatorial Aspect of Anticipants

Guided by the self-similarity, not-yet-identified visibles $\Xi$ efficiently scan chromatic complexity of object images. Following the fractal sampling, the complexity of brightness distribution is captured in terms of trichromatic components $R_\omega$, $G_\omega$, $B_\omega$. The trichromatic decomposition is essentially mental process; for most human observers only three primers are required to match a test light. On such cognitive triple $(R, G, B)$, define

$$\phi(f_\omega) = \frac{f_\omega}{|f_\omega|}, \quad |f_\omega| = \sqrt{R^2_\omega + G^2_\omega + B^2_\omega}.$$

With this nonnegative random variable $\phi_\omega = \phi(f_\omega)$, we can introduce the following geometric representation, called chromatic information space:

$$\Phi = \{ \phi(f) \mid |\phi(f)| = 1 \}.$$ \quad (4)
Define local metric in $\Phi$ by

$$g_\alpha(\phi_i|\phi_j) = \exp\left[-\frac{|\phi_i - \phi_j|^2}{2\alpha}\right].$$

(5)

The resolution parameter $\alpha$ is adjusted to the range of latent whiteness in scene imagery [8]: *An illuminant may be up to 93% chromatic, but provided it it contains at least 7% "daylight", surfaces with uniform spectral reflectance – that reflect equally at all wavelength – will remains achromatic.*

Consider the following set $s \in \Phi$ associated with the fractal sampling $\Xi$:

$$s = \{ \phi(\xi) \in \Phi \mid \xi \in \Xi \}.$$  

(6)

With such a palette, we have the following conditional probability as chromatic aspect of the anticipants:

$$dP(\omega|s) \propto \gamma_\alpha(\omega|s)dP(\omega),$$

(7)

where $\gamma_\alpha(\omega|s) = \max_{\phi_j \in s} g_\alpha(\phi(\omega)|\phi_j)$.

### 7 Control Aspect of Anticipants

Through the computational and chromatic aspects, the anticipants are visualized as shown in Fig. 13. Via statistical analysis on $\varphi_\rho(\omega|\nu)$, vision system estimates the expansion of roadway to extract the vector $\mathbf{v}$ in terms of a segment on scene image.

Let the totality of possible maneuvering can be identified with stochastic process $X$ along the vector $\mathbf{v}$. By mapping time parameter $t$ on $\mathbf{v}$, possible the trajectory of the vehicle is represented as a sample path $x_t$, $t \in \mathbf{v}$ on the aspects. Within such stochastic framework, the maneuvering process governed by
human’s *in situ* intention is represented as shown in Fig. 14. By coding the boundaries in $\varphi\nu$-aspect, we have a pair of boundary vectors $(\nu_+, \nu_-)$ given by

$$v_\pm \sim \frac{\varphi_\nu(\xi_+|\nu)}{\varphi_\nu(\xi_+|\nu) + \varphi_\nu(\xi_-|\nu)},$$  

(8)

Suppose that physical maneuvering process is observed by $y_\tau$. By mapping physical time $\tau$ on $v$, we have the information $F^\nu_t$ consisting of $y_s$, $s \leq t$. Regarding the information $F^\nu_t$, the class $X$ is divided into the past and the future in $\varphi\nu$-aspect. This induce two sub-areas in the aspects of anticipants as illustrated in Fig. 14; in the horizon of control, sample path of trajectory $x_t$ is determined with respect to $F^\nu_t$; in the scope of the perception, the trajectory $x_t$ is generated via ‘agreeable’ human-machine interaction.

8 Square Root LV-System

The trajectory in the scope of the perception is governed by essentially spontaneous decision of humans. Simultaneously, only agreeable trajectories should be included in the anticipants. To satisfy these requirement, the trajectories should be confined by the boundaries $(\nu_+, \nu_-)$ throughout the to time evolution. Consider the the following Lotka-Volterra system with respect to squares of the populations $(x^2, y^2)$:

$$\frac{1}{x^2} \frac{dx^2}{dt} = \alpha(x^2 + \sigma^2 y^2),$$  

(9a)

$$\frac{1}{y^2} \frac{dy^2}{dt} = -\alpha(y^2 + \sigma^2 x^2),$$  

(9b)

with $\alpha(\cdot) = 1 - \kappa(\cdot)$. Noticing that the system confines the solution $(x^2, y^2)$ within the space of positive populations even under noisy disturbances and observation error [17], let the positive (negative) part of the process $x_t$ be identified with $x_+$ $(x_-)$ satisfying the following sub-processes:

$$2 \frac{1}{x_+} \frac{dx_+}{dt} = 2 \frac{1}{x_-} \frac{dx_-}{dt} = \alpha(x^2 + \sigma^2 y^2).$$

Suppose that $(x_+, x_-)$ sub-process are observed via the following channels, respectively:

$$2 \frac{1}{y_+} \frac{dy_+}{dt} = -\alpha(x^2_+), \quad 2 \frac{1}{y_-} \frac{dy_-}{dt} = -\alpha(x^2_-).$$

Define $(\cdot)_+ = (\cdot)_+ - (\cdot)_-$. Since

$$2dy_+ = [\tilde{\kappa}x^+_+ - y^+_+] dt + \sigma dw,$$

with $\tilde{\kappa} = \sqrt{x^2 y^2/\kappa}$, under human’s spontaneous decision $w_t$ and noticing the Volterra’s principle $x_+ y_+ \sim x_- y_- \sim 1/\kappa$, it follows that

$$\frac{dx_+}{x_+} = \frac{1}{2} \alpha \left( x^2_+ + \sigma^2 (y^+_+)^2 \right) dt + \frac{1}{2} R_+ dv_+,$$

(10a)

$$\frac{dy_+}{y_+} = \frac{1}{2} \left[ x^+_+ - y^+_+ \right] dt + \frac{1}{2} \sigma_t dw.$$

(10b)

where $(v_+, v_-)$ and $w$ stand for random excitations and observation error, respectively. In this system, $(v_+, v_-)$ denotes the uncertainty of boundary code $(\nu_+, \nu_-)$.

The schematics of the multi-aspect representation of anticipants is illustrated in Fig. 15 where physically future trajectory $x_t$ is generated as a stochastic process under the feedback of humans ‘agreeable’-but-spontaneous decision $y_t$. The introduction of jointly ‘white Gaussian’ models $(v, w)$ implies that the anticipants describe future trajectory without any restriction on unpredictability of scene and resulted human’s voluntary operation. By the nonlinear interaction of joint $(x_t, y_t)$-process, the uncertainty of boundaries $(\nu_+, \nu_-)$ are autonomously recovered in cooperation with human’s possible operation observed via $y_t$ as shown in Fig. 16; even in the case of sudden loss of boundary information $(x_+, x_-)$ at ‘stopping time’, possible human’s avoidance operation is generated as $y_t$ to restore the boundaries stochastically.

The implication of stochastic cooperation processes is investigated through simulation studies. The situation to be simulated is illustrated in Fig. 17; the problem is to anticipatively evaluate the trajectory to be generated in real scene in segmented roadway model. Figure 18 indicates the trajectory generated on segmented roadway model for possible following to real pattern; in subsequent straight roadway, the deviation of $y_t$ rapidly converges to the equilibrium as demonstrated in Fig. 19.
For anticipative following to the agreeable path as shown in Fig. 17, the cooperative control process invokes \textit{a priori} roadway model segmented in terms of $v$. To extract a chain of the segments in \textit{as-is} imagery, on the other hand, a “road following process” is applied to the combinatorial aspect of the anticipants as shown in Fig. 20. This implies that the anticipants yields theoretical basis for \textit{in-situ}
Multi-Aspect Approach to Robotic Anticipation for Supporting Human’s Multiple-Intelligence in Natural Scene

9 Discussions

For robotic systems, one of essential difficulties arises in paradoxical process supporting human’s spontaneity: pre-experienced perception followed by preestablished decision. Via setting-up of informatic vicinity on vehicle-satellite-roadway network, human-robot interaction is confined in a restricted anticipation through with computational, combinatorial and control aspects on scene image.

On the anticipants, decision processes are represented in terms of stable dynamics: iterated function system for complex pattern imaging and stochastic differential equation for agreeable trajectory generation. Despite formal separation, the dynamical systems are intrinsically coordinated through concurrent access to the anticipants. Mathematically, the coherence without explicit interaction is maintained by statistical identity of the ‘whiteness’ introduced as general representation of uncertainty in each aspect. By generating and indicating such anticipants, open logic machine can substitute a part of mental processes as shown in Fig. 21 where robotic intelligence should be accepted as a part of as-is scene image.

Multi-aspect approach, as displayed in Fig. 9, is founded on something-invariant throughout a session consisting of precautious perception and spontaneous decision. Despite inherent latency, the something-invariant generates aspect specific images as shown in Fig. 13. The visibility of the multi-aspect approach can be extended to generalized human-robot interaction as shown in Fig. 22: being supported by well-
organized internal dynamics, the consistency of humans performance should be observed as something-invariant; by identifying the control aspect and the computational aspect with the space of stable signals and the totality of decidable processes, respectively, adaptation-reasoning scheme are regulated partially in each aspect towards the latent consistency; in situ restriction of object space as the combinatorial aspect of the mental space confine the range of the adaptation-decision processes within the interior of the aspects. Such asynchronous untangle nondeterministically restore neuro-physiological loss in the critical one second in cooperation with the mental expansion observed as Fig. 8.

10 Concluding Remarks

The concept of anticipants has been introduced for tightening human-robot interaction in information intensive maneuvering. Through multi aspects of the anticipants, robotic intelligence is co-oriented with human’s spontaneous decision. The implementation of the anticipants on network system is the next problem.

References