Morphology-Based Object Coding: Self-Similarity Detection via Pattern Sensitive Image Sampling

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

Abstract: Being intended for reactive access to open information space, a method is presented of object description in noisy background. The description is transparent so that visible information is contoured and parametrized without any a priori information. For this purpose, first, visibles are approximated by fractal attractors, without serious loss of generality. By evaluating capturing probability of unknown fractal attractors, next, most probable points are extracted to restore self-similarity structure to be identified. By using proposed method, finally, computer vision system is shown to be equipped with the capability for “picking up” not-yet-identified objects as robots can interact with unknown objects under complete control of human operators.

Keywords: Object Coding; Morphological Features; Fractal Collage; Self-Similarity Detection; Pattern Sensitive Sampling

Table of Contents
1 Introductory Remarks 1
2 Reactive Knowledge Space 2
3 Fractal Computability of Random Patterns 2
4 Machine Perception of Self-Similarity 4
5 Generic Representation of Self-Similarity 7
6 Discussions 9
7 Concluding Remarks 9

1Proceedings of the 10th IEEE International Workshop on Robot and Human Interactive Communication (RoMan01), September 18–21, 2001, Paris-Bordeaux, France
1 Introductory Remarks

Being intended for industrial applications, conventional robotics deals with the adaptation to nonanticipative environment as a version of fundamental “frame” problem. Inside real world, however, the setup for knowledge integration is rather simple. By restricting knowledge within ‘common pool’, the performance of individuals is \textit{a priori} acceptable by co-existents of a fixed community \cite{10}. By letting real world to select authentic knowledge of this sort, decision processes are \textit{a posteriori} maintained within consistent solution spaces. Such nondeterministic knowledge handling scheme has been invoked to formulate linguistic communication model \cite{4} and interactive computing system, called \textit{open logic machine} \cite{12}, as well.

Open logic machines successively generate selective representation of programmed knowledge on perceptual object code. By restricting the range of the programmable knowledge through perception, hence, we can simulate the knowledge integration process on computing systems. As constituents of nature, humans seem to adapt their knowledge to real world without any difficulty. The adaptation process is so subtle that scenes reveal themselves as if their appearances can afford to provide definite cue for consistent decisions \cite{5}. Sophisticated adaptation processes of that sort should be be supported by intentional perception channel to organize stimuli into structured affordance specifying the scene \cite{14}. Computational decision on knowledge often generate possibility scenes that is inconsistent with real world. This implies that direct coupling between conceptual knowledge and existing environment should rather restrict computation process than complete stored knowledge.

2 Reactive Knowledge Space

Knowledge restriction via perception not only provides computational basis of individual intelligent performances but also completes communication channel for cooperative decision making. Noticing architectural equivalence, knowledge space restricted by on open logic machines can be integrated with knowledge pool through environment sharing. Within a specific scene interconnected with humans and open logic machines, in other words, human’s scope of knowledge integration can be extended to networked community. In fact, various kinds of knowledge can easily be transferred to humans in specific situations via \textit{imitation} process \cite{3}. In the imitation process, extended classes of information including tunes, ideas, catch-phrases, clothes, fashions, ways of making pot or of building arches, are believed to be “encoded” as \textit{meme} and, then, propagated by leaping from brain to brain. This implies transferred knowledge should be encoded in terms of receptor specific syntax representing the reproduction of the scene.

Computing systems, on the other hand, transfer digital files without any loss nor restriction as their natures. Via the open logic access to computer network, hence, humans seamlessly expand their horizon of thought to entire digital knowledge sources. This implies that, insofar as human’s reference use, today’s computer network realized the dream of early computer scientists \cite{2}. Furthermore, the establishment of software platforms, digitized knowledge can easily be instilled into user specific systems. By the completeness of formal system within closed world of “pure-logic”, however, there is no way to maintain knowledge handling processes consistent with external real world. Intelligent robots are adapted to external world only through environment model. Direct coupling via multiple feedback mechanisms exhibits physical reactions that can be represented as significant behaviors only through the empathy of human observer. Such separation of programmed knowledge from real world yields serious difficulty in the \textit{a posteriori} stabilization of interactively instilled knowledge system.

The difficulty for computer implementation of the human-environment coupling arises from contradictive perception process itself: \textit{to generate object symbols for specifying scene as the references of scene analysis}. Real world can afford sufficient cues for all participants intended for their own goals. This implies that affordances are not always convey significant messages for specific participant. Hence, each decision maker is requires to extract unpredictable objects in noisy background image (Fig. 1). \cite{8} Recent advancements in mathematical vision theory reveals that essential process of visual perception can be modeled by Markov processes on rather simple hardware structure \cite{15}. It has been pointed out that wiring structure of layered...
Referent objects of complex appearances are observed with many “distracting” objects. To concentrate objects of interest, perception channels are required to generate symbols of unpredictable objects in “background noise.”

Mathematical structure, the self-similarity, governs the object of interest. Abstract description is not localized so that the object can be “pick up” by contouring random distribution of “domain features.”
For any pattern $\Lambda$ in a fixed image plane $\Omega$, there exists a set of contraction mappings $\nu = \{\mu_i\}$, $\mu_i : \Omega \rightarrow \Omega$ that yields an invariant subset $\Xi \subset \Omega$, approximating the pattern $\Lambda$, within arbitrary small imaging error.

Retina should be considered as the generator of multi-scale Gaussian distribution [7]. The introduction of Markov process on Gaussian fields implies a stochastic channel to generate primal sketch for supporting subsequent recognition processes [13] or feature pattern for designing imaging processes [9].

3 Fractal Computability of Random Patterns

A potential way to bypass the self-contradiction of perception channel is to introduce the self-similarity as a priori pattern structure. Noticing logical-geometric coordination in self-similarity imaging processes, in this paper, we assume that patterns to be observed are generated as fractal attractors associated with unknown set of contraction mappings. Without serious loss of generality, in the following discussions, we suppose that the number of contraction mappings can be guessed. The assumption of self-similarity is not so restrictive because we can approximate any patterns in terms of the following “Fractal Collage [1]”: For any pattern there exists a set of contraction mappings that yields an invariant subset for approximating the pattern within arbitrary small imaging error. This implies that any observed patterns can be coded in terms of finite symbols. The finite code specifies imaging process for generating fractal attractor of infinite geometric complexity as pattern model.

In contrast with conventional statistical-computational representation, fractal model conveys complete information to specifies invoked contraction mappings. In fact, we have enough data for determining mapping parameter as the distribution of attractor points. Hence, we have logical bases for pattern coding as the following “Structural Observability [6]”: The attractor is covered by the totality of fixed points associated with all finite composite of the mappings. Thus, pattern coding results in identifying origin-destination pairs in complex attractors. Since each attractor point is deterministically “jumped” by a contraction mapping, we have exact origin-destination associations in observed imagery.

In addition, the self-similarity induces definite association between geometric order, i.e., spatial distribution of attractor points, and probability for pattern capturing, i.e., gray level distribution. This implies
that we can analyse pattern structure via the estimation of the “Invariant Measure [1]”: For arbitrary attractor generated by random application of fixed contraction mappings, there exists a measure that is invariant with respect to transform by the mappings. The existence of invariant measure implies the association between the distribution pattern and the density function of fractal attractors. The self-similarity of the density function introduces the self-similarity in the distribution of statistical parameter.

We can exploit the collage theorem as a general framework for fractal pattern coding. Fractal code is described in terms of finite contraction mappings that restore observed patterns of infinite complexity. By invoking the structural observability, we can design the mappings through origin–destination association on discrete points. To discriminate the discrete image from background noise, the invariance of observed “brightness distribution” with respect to the mappings to be designed should be analyzed. In this paper, hence, we consider the integration of these three aspects of the self-similarity to develop a unified sampling scheme for unknown complex patterns. From the viewpoint of practical applications, in what follows, the sampling scheme is designed to require only the size of mapping set as a priori information.

4 Machine Perception of Self-Similarity

Let $\Omega \subset \mathbb{R}^2$ be a continuous image plane and suppose that patterns are generated within the Borel field $\mathcal{F}[\Omega]$ of the totality of subsets of $\Omega$. The disparity between patterns in $\mathcal{F}[\Omega]$ is indexed in terms of the Hausdorff distance $\eta[\cdot, \cdot]$ defined by

$$\eta[A, B] = \max \{\eta[A, B], \eta[B, A]\}, \quad (1a)$$

$$\eta[A, B] = \max_{\omega \in A} \{\min_{\lambda \in B} |\omega - \lambda|\}, \quad (1b)$$

for $A, B \in \mathcal{F}[\Omega]$. Consider a fixed set of unknown contraction mappings $\nu = \{\mu_i, i = 1, 2, \ldots, m\}$ with length $\|\nu\| = m$, where $\mu_i : \Omega \rightarrow \Omega$ is a mapping from $\Omega$ into itself with contractivity factor $s_{\mu_i}$, $0 < s_{\mu_i} < 1$, i.e.,

$$|\mu_i(\omega_1) - \mu_i(\omega_2)| \leq s_{\mu_i}|\omega_1 - \omega_2|, \quad (2)$$

for any $\omega_1, \omega_2 \in \Omega$. By the contractivity, we can program pattern generation processes as the “collage” of mapping images $\mu_i(\Omega)$. For self-similar patterns, particularly, we have the following exact pattern generation scheme on program set $\nu$:

**Proposition 1** (Fractal Attractor) Let $\Xi$ be the attractor generated by the “program” $\nu$ to satisfy

$$\Xi = \bigcup_{\mu_i \in \nu} \mu_i(\Xi). \quad (3)$$

The attractor $\Xi$ can be successively approximated by the dynamical system on $\mathcal{F}[\Omega]$:

$$\xi_{t+1} \in \bigcup_{\mu_i \in \nu} \mu_i(\Xi_t), \quad (4a)$$

$$\Xi_t = \{\xi_\tau \in \Omega, \tau \leq t\}. \quad (4b)$$

The sequence $\Xi_t$ converges to the attractor $\Xi$ in the following sense

$$\lim_{t \to \infty} \eta[\Xi_t, \Xi] = 0. \quad (5)$$
By assigning the following basic measure to the image plane:

\[ dP(\omega) = \frac{d\omega}{\int_{\Omega} d\omega}, \tag{6} \]

we can introduce a probability space \((\Omega, \mathcal{F}[\Omega], P)\) as the basis of image analysis. For instance, the “brightness” of pattern \(\Xi\) can be evaluated by

\[ f^\xi_\sigma * \chi_\Xi(\omega) = \lim_{t \to \infty} \frac{1}{\|\Xi_t\|} \sum_{\xi \in \Xi_t} f^\xi_\sigma(\omega), \tag{7a} \]

where \(f^\xi_\sigma\) is uniform test function at \(\xi\):

\[ f^\xi_\sigma(\omega) = \begin{cases} \frac{1}{2\pi\sigma^2}; & \text{for } |\omega - \xi| \leq \sigma, \\ 0; & \text{otherwise}. \end{cases} \tag{7b} \]

For fixed \(\xi\), the value \(f^\xi_\sigma * \chi_\Xi(\omega)\) provides the number of point images at \(\omega\). Let \(\xi\) moves on the following lattice \(\mathcal{L}\):

\[
\begin{bmatrix}
\cdots & (i - \epsilon, j + \epsilon) & (i, j + \epsilon) & (i + \epsilon, j + \epsilon) & \cdots \\
\cdots & (i - \epsilon, j) & (i, j) & (i + \epsilon, j) & \cdots \\
\cdots & (i - \epsilon, j - \epsilon) & (i, j - \epsilon) & (i + \epsilon, j - \epsilon) & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots
\end{bmatrix}
\]

with \(\epsilon = \sigma/\sqrt{2}\). By using a system of test functions \(\{f^\ell_\sigma, \ell \in \mathcal{L}\}\), we have the gray level image of fractal attractor \(\Xi\) as shown in Fig. ??, where point images with distribution \(\chi_\Xi\) are located on image plane with uniform background noise with distribution \(\chi_\Omega\). In this case, the intensity of noise \(\|\Omega\|\) is adjust to four times of the pattern intensity, i.e., \(4 \times \|\Xi\|\).

For generating smooth distribution, consider a system of Gaussian distributions given by

\[ \mathcal{G} = \{ g^\ell_\sigma, \ell \in \mathcal{L} \}, \tag{8a} \]

\[ g^\ell_\sigma(\omega) = \frac{1}{2\pi\sigma^2} \exp \left[ -\frac{|\omega - \ell|^2}{2\sigma^2} \right]. \tag{8b} \]

By testing the value of distributions on \(\mathcal{G}\), we have the following stochastically sampled image:

\[ \chi_\Xi^\mathcal{G} = \{ g^\ell_\sigma * \chi_\Xi(\omega), \ell \in \mathcal{L} \}. \tag{9} \]

In the sampling scheme \((\mathcal{G}, \mathcal{L})\), complete information \(\chi_\Xi\) of infinite resolution is associated with discrete image plane \(\mathcal{L}\). For instance, unknown self-similarity mappings yields the following “counting rule” for input filtering pattern area on digital image plane \(\mathcal{L}\):

\[ P(N^\Xi_\nu) \geq \|\nu\|, \tag{10} \]

where \(N^\Xi_\nu\) means the number of point images located in the neighborhood of the Gaussian array \(g^\ell_\sigma \in \mathcal{G}\).

The imaging process (4) expands initial points \(\Xi_0\) through nondeterministic scattering within a fixed domain \(\Xi\). This implies that the process should be modeled by 2D dynamical system with the following antagonistic imaging mechanisms

- diffusion of point image at \(\xi\) within image plane \(\Omega\), and,
Point images on $\Xi$ are located on an image plane $\Omega$ with uniform background noise. The number of background noise of intensity $\|\Omega\|$ is adjusted to four times of pattern intensity $\|\Xi\|$ evaluated in terms of the number of point images in $\Xi$.

- successive reduction of imaging domain via not-yet-identified contraction mappings $\mu_i \in \nu$.

By representing this imaging mechanism on sampled image $(\Theta, \Sigma)$, it follows that

**Proposition 2** (Capturing Probability) Let $\chi_\Xi$ be a given brightness distribution to be collaged by $\nu = \{\mu_i\}$. Assume that the distribution is observed through the Gaussian array $\Theta$. Then the probability for regenerating $\Xi$ within the framework of maximum entropy capturing is visualized as a smooth field $\varphi(\omega|\nu)$ satisfying, for each $\ell \in \mathcal{L}$,

$$\frac{1}{2}\Delta \varphi(\omega|\nu) + \rho |g^\ell \ast \chi_\Xi(\omega) - \varphi(\omega|\nu)| = 0,$$

where $\rho = \log \|\nu\|$ is the complexity parameter.

The capturing probability $\varphi(\omega|\nu)$ is visualized as smooth gray level image as shown in Fig. 5. On smooth field $\varphi(\omega|\nu)$, we can easily extract discrete pattern

$$\tilde{\Theta} = \left\{ \hat{\theta} \in \Omega \mid \nabla \varphi = 0, \det [\nabla \nabla^T \varphi] > 0, \Delta \varphi < 0 \right\},$$

through locally parallel image analysis. By definition, discrete pattern $\tilde{\Theta}$ is a sampling of fractal attractor $\Xi$ stochastically dependent on imaging parameter $\nu$. The dependence is implicit: only mapping size is required as a priori information.

5 Generic Representation of Self-Similarity

We can generate the capturing probability for fractal patterns in background noise as shown in Fig. ???. In background noise, however, differential operators $\nabla$, $\Delta$ yield local maxima exterior of the patterns. For contouring in sampled image $\Theta$, define a version of conditional probability for evaluating possible variation
The cross section of the solution to stationary diffusion equation (11) with $\rho = 4$ and $\epsilon = 0.2$ is indicated. Thus, $\varphi(\omega|\nu)$ visualizes the stochastic evaluation of a point image emitted from the origin as a smooth field.

of brightness, as follows:

$$p(\omega|\nu) = \frac{\varphi(\omega|\nu)}{C^\varphi_\Omega},$$  \hfill (13a)  

$$C^\varphi_\Omega = \int_\Omega \varphi(\omega|\nu) d\omega.$$  \hfill (13b)

By using $p(\omega|\nu)$, we can index the complexity of brightness variation in terms of the following Shannon’s entropy

$$\hat{H}_\nu = -\mathbb{E}\left\{ \log p(\omega|\nu) \mid \nu \right\}.$$  \hfill (14)

The existence of self-similarity structure should be verified through the comparison with the entropy evaluation under “null condition”:

$$\hat{H}_\emptyset = -\mathbb{E}\left\{ \log p(\omega|\emptyset) \mid \emptyset \right\},$$  \hfill (15)

where $p(\omega|\emptyset) = \text{const}$. By the comparison of complexity indices, $\hat{H}_\nu, \hat{H}_\emptyset$, we have

**Proposition 3** (Output Filter, Fig. 6) [11] Assume the background noise $\chi_\Omega$ is uniformly distributed in the image plane $\Omega$ and suppose that observed measure $\chi_\Lambda$ is represented by

$$\chi_\Lambda = \chi_\Xi + \chi_\Omega.$$  \hfill (16)

Suppose that $p(\hat{\theta}|\nu) \geq \bar{p}_\nu$ where

$$\log \bar{p}_\nu = 1 - \frac{1}{2}(1 - e^{\hat{H}_\nu - \hat{H}_\emptyset}) - \hat{H}_\emptyset.$$  \hfill (17)

Then $\hat{\theta} \in \Xi$.  

---

*Figure 5: Multi-Scale Image*

The cross section of the solution to stationary diffusion equation (11) with $\rho = 4$ and $\epsilon = 0.2$ is indicated. Thus, $\varphi(\omega|\nu)$ visualizes the stochastic evaluation of a point image emitted from the origin as a smooth field.
Define
\[ \hat{\Theta} = \left\{ \hat{\theta} \in \tilde{\Theta} \mid p(\hat{\theta} | \nu) \geq \bar{p}_\nu \right\}. \]

This filtered sampling image is well structured in the following sense:

**Proposition 4 (Invariant Features)** Assume that there exists a subset \( \Theta \subset \hat{\Theta} \) invariant with respect to \( \nu \), i.e.,
\[ \Theta = \left\{ \theta \in \hat{\Theta} \mid \exists \mu_i \in \nu : \mu_i^{-1}(\theta) \in \Theta \right\}. \quad (18) \]

Suppose that for arbitrary \( \mu_i \in \nu \) there exist \( \theta^o, \theta^d \in \Theta \) such that
\[ \theta^d = \mu_i(\theta^o). \quad (19) \]

Then for arbitrary attractor point \( \xi \in \Xi \), there exists a finite sequence \( \cdots \mu_k \mu_j \mu_i \) that yields a fixed point to approximate \( \xi \) within a given accuracy.

By applying mapping sets associated with dictionary patterns, we can detect contraction mappings as shown in Fig. 7. In this case, the observability test (19) select true pattern successfully.

6 Discussions

As shown in equations (11), (12) and (17), we can extract a discrete information \( \hat{\Theta} \) as a transferable knowledge. In fact, \( \Theta \) is generic information of unknown pattern to support computable specification (18). Noticing that the extraction process is essentially dependent on only “class” parameter \( ||\nu|| \), furthermore, we can preset perception channel in terms of a priori guesses of \( ||\nu|| \). Many results in fractal graphics implies that various interesting images can be generated via self-similarity processes with \( ||\nu|| = 3 \) or 4. Despite nondeterministic formulation (18), we can design mapping set through cut-and-try on reduced search space \( \hat{\Theta} \).

7 Concluding Remarks

Finite feature based fractal coding is considered within the context of reactive access to open information space. Based on the capturing probability of unknown fractal attractors, unknown fractal attractors are
The selection of mappings satisfying the origin–destination associations (19) on filtered pattern \( \Theta \) yields observable code. Through such observability analyses, hence, nearmiss patterns were eliminated.

Sampling into finite feature set. The structural consistency of the finite features was demonstrated to yield efficient detection scheme for fractal code. Despite serious contouring error in observed imagery, the consistency of features were shown to be maintained.

References