FRACTAL CODING OF IMAGE-COLOR SPACES FOR SALIENCY-BASED OBJECT DETECTION IN NATURALLY COMPLEX SCENES

Kohji Kamejima

Faculty of Information Science and Technology, Osaka Institute of Technology
1-79-1 Kitayama, Hirakata 573-0196 JAPAN
phone: +81-72-866-5406, fax: +81-72-866-8499, email: kamejima@is.oit.ac.jp
web: http://www.is.oit.ac.jp/~kamejima

ABSTRACT

A saliency-based approach is presented for object detection in naturally complex scenes. By regenerating the chromatic diversity in a probabilistic color space, the distribution of saliency colors is extracted as the viewer specific visualization of landmark objects. The saliency distribution is articulated into a system of fractal attractors spanning object images. Detected fractal models are visualized according to the perspective underlying the scene image.

1. INTRODUCTORY REMARKS

Despite infinite diversity of appearance, natural scenes exhibit environment specific landmarks to be identified within individual intention of viewers. To control the focus to such a landmark object, perception processes should gather randomly distributed image features and apply 'feature integration' schemes to generate 'visual saliency' associated with the complex scene [3]. Due to the redundancy of natural scene relative to specific decision making by the multitude of the viewers, however, such computational feature integration processes easily fall into combinatorial explosion.

As the results of the evolution in the really existing world, human's vision system is equipped with not-yet-explicated information processing mechanism for understanding the scenes thronged with friendly or undesirable neighbors. Through the co-evolution process, the inherent vision has developed an attention control mechanism within the surroundings [5] on the premise that imminent decision making should be evoked by early perception of unstructured ambient light [1] and universal preference to a class of fractal patterns [2]. Based on such physical-geometric structure underlying the naturally complex scenes, we can articulate the scene image into two fractal attractors spanning a connected open space and a distribution of boundary objects as shown in Fig. 1; the ground area and boundary distribution are visualized by a closed link and capturing probability with respect to associated fractal models, respectively [4].

Noting this, in this paper, we consider the multi-fractal coding of the boundary distribution in scene image. The problem is to extend the fractal articulation to the chromatic diversity arising in a color space.

2. LOCALLY GAUSSIAN PALETTE

To cooperate with human's inherent perception, the chromatic diversity of the boundary distribution should be regenerated in terms of at least three 'primary colors'. Let Ω be a image plane and suppose that the information conveyed by the ambient light is observed through a distribution of incoming light f_ω on Ω. For precise evaluation of the chromatic complexity, let the saliency of the incoming light be detectable through spectral factorization: $f_\omega^{\mathrm{RGB}} = \left[f_\omega^{\mathrm{R}} \quad f_\omega^{\mathrm{G}} \quad f_\omega^{\mathrm{B}}\right]^T$, $\omega \in \Omega$, where $f_\omega^{(\cdot)}$ designates subjective weight of the primary $(\cdot)$. Define $\phi_\omega = f_\omega^{\mathrm{RGB}}/|f_\omega^{\mathrm{RGB}}|$. By identifying the totality of the chromatic information ϕ_ω with the positive part of a unit sphere Φ^+ we can induce the following index:

$$g_\alpha(\phi|\phi_\omega) = \frac{1}{2\pi\alpha}\exp\left[-\frac{|\phi-\phi_\omega|^2}{2\alpha}\right], \qquad (1)$$

for arbitrary $\phi \in \Phi^+$; following experimental studies using various roadway scene images, the sensivity factor α should be adjusted to $1/10 \sim 1/100$. For sufficiently small chromatic variation, the index $g_\alpha(\phi|\phi_\omega)$ approximates the Gaussian distribution on local tangential space at ϕ_ω. Adding to the local measure, the constraint $|\phi_\omega|^2 = 1$ yields the following index for evaluating the predictability of the chromatic impression:

$$\psi_\omega = \exp[-\mathcal{H}_\omega], \qquad (2)$$
$$\mathcal{H}_\omega = -2\sum_{\mathrm{RGB}}\left(\frac{f_\omega^{(\cdot)}}{|f^{\mathrm{RGB}}|}\right)^2 \log\left(\frac{f_\omega^{(\cdot)}}{|f_\omega^{\mathrm{RGB}}|}\right).$$

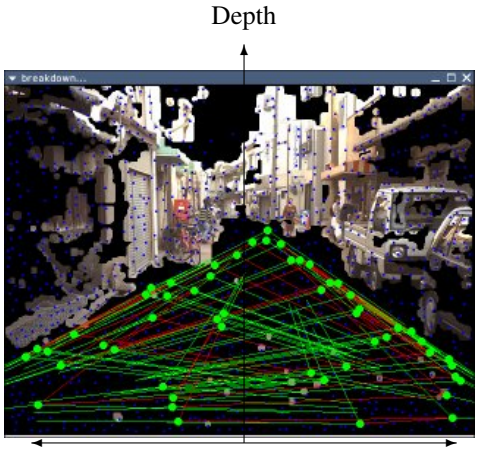
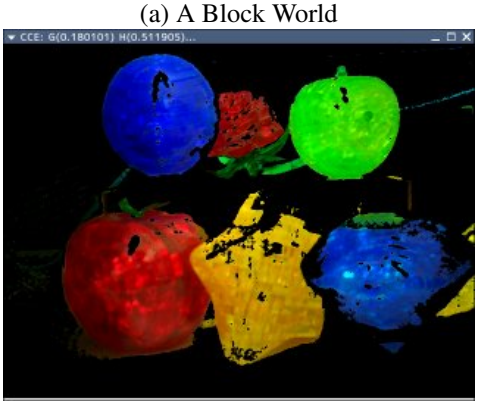


Figure 1: Generic Structure of Naturally Complex Scene
a connected open space is identified with a fractal code spanning scale features (green circles); inherent perception articulate the boundary distribution into landmarks exhibiting saliency colors.

(a) A Block World



(b) Filtered Image

Figure 2: Complexity Reduction via ψ_ω Channel
an artificial scene can be articulated with respect to conventional trichromatic primary.

Figure 3: Complexity Reduction via ψ_ω Channel
trichromatic articulation may miss and/or degenerate some saliency patterns (?) in naturally complex scenes.

In this indexing, the complexity of substantial process implemented by the retina system is evaluated in terms of Shannon's entropy $\mathcal{H}_\omega^{\mathrm{RGB}}$; the complexity reduction in subsequent mental process is identified with vitals specific 'neg-entropy' generation. By using the indexing (2), we can visualize saliency distribution in well-structured scene as illustrated in Fig. 2 where the chromatic diversity of the 'block world' (a) is 'matted' with respect to the index ψ_ω to yield the saliency patterns in noisy background (b).

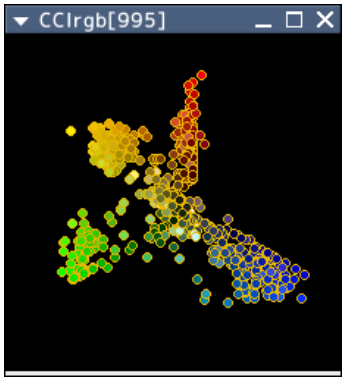


Figure 4: Chromatic Diversity in RGB Color Space
the chromatic diversity associated with artificial objects yields 2D distribution essentially expanded towards the trichromatic primary.

The saliency indexing (2) can be applied to naturally complex scene to generate the distribution $f_\omega^{\mathrm{RGB}}\psi_\omega$ as shown in Fig 3 where landmark objects distributed in the scene (Fig. 1) are partially visualized; objects exhibiting 'natural colors' may pass through the filtering process. To correct such substantial bias of the ψ_ω-filter, we need the adaptation of the primary system to the chromatic diversity of the scene.

3. CHROMATIC COMPLEXITY GENERATOR

Consider the fractal model generating the chromatic complexity as shown in Fig. 4. In this figure, the diversity of the chromatic information in the scene (Fig. 2) is visualized in the following linear color space:

$$\begin{aligned} \Gamma \ni \gamma \;\; = \;\; e^{\mathrm{RGB}}\phi_\omega, \qquad e^{\mathrm{RGB}} = \begin{bmatrix} e^{\mathrm{R}} & e^{\mathrm{G}} & e^{\mathrm{B}} \end{bmatrix}, \\ e^{(\cdot)} = \begin{bmatrix} \cos\theta_{(\cdot)} & \sin\theta_{(\cdot)} \end{bmatrix}^T, \end{aligned} \tag{3}$$

with *a priori* orientation of the trichromatic primaries $\theta_{\mathrm{R}} = \pi/2, \theta_{\mathrm{G\,(B)}} = \theta_{\mathrm{R}} + (-)2\pi/3$; the distribution of the chromatic information is expanded towards the preassigned RGB primaries in the color space Γ. This implies that the chromatic complexity of the natural scenes, partly missed in Fig. 3, can be regenerated through the adaptation of the primary to the observed images. Let $\hat{\Pi} = \{\hat{\pi}_i\}$ be a set of such *as-is* primaries with size $\|\hat{\Pi}\|$. Suppose that samples of the chromatic information $\mathfrak{s}$ are collected in a scene image and diffused via the following process in Γ:

$$\frac{\partial \varphi_\rho(\gamma|\mathfrak{s})}{\partial t} \;\; = \;\; \frac{1}{2}\Delta\varphi_\rho(\gamma|\mathfrak{s}) + \rho\left[\chi_\mathfrak{s} - \varphi_\rho(\gamma|\mathfrak{s})\right], \tag{4}$$

where $\chi_\mathfrak{s}$ denotes the aggregation of Dirac's delta measure distributed on the set $\left\{\gamma(\phi_\omega) \;\middle|\; \phi_\omega \in \mathfrak{s}\right\}$. By adjusting $\rho = \log_2\|\hat{\Pi}\|$, we can identify the information $\varphi_\rho(\gamma|\mathfrak{s})$ with the probability distribution for capturing a fractal attractor controlled by a set of fixed point. Hence, we have the following procedure for an estimate of the *as-is* primary $\hat{\Pi}$ on the Laplacian-Gaussian boundary $\partial^g\chi_\mathfrak{s}$ [4].

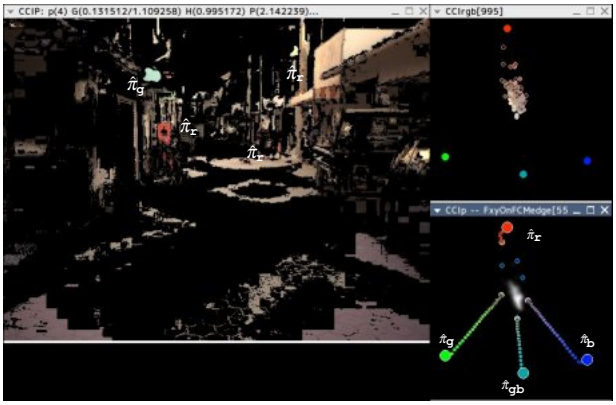


Figure 5: Complexity Reduction via $\hat{\psi}_\omega$ Channel
the chromatic diversity of a naturally complex scene is reduced into saliency patterns with respect to shifted as-is primaries $\hat{\pi}_\text{r}$ and $\hat{\pi}_\text{g}$.

(i) Fixed Point Allocation: First, a boundary point $\tilde{\partial}\gamma^f \in \partial^g\chi_\mathfrak{s}$ is expanded via the following successive scheme:

$$\begin{aligned}\tilde{\Gamma}^f_{t+1} &= \tilde{\Gamma}^f_t \cup d\tilde{\Gamma}^f_t, \qquad (5)\\ d\tilde{\Gamma}^f_t &= \Big\{\tilde{\partial}\gamma^f \;\big|\; \forall\partial\tilde{\gamma}:\overleftarrow{\eta}\left(\partial\tilde{\gamma}^f,\tilde{\Gamma}^f_t\right)\geq\overleftarrow{\eta}\left(\partial\tilde{\gamma},\tilde{\Gamma}^f_t\right)\Big\},\\ &\qquad \tilde{\partial}\gamma^f,\partial\tilde{\gamma}\in\partial^g\chi_\mathfrak{s}-\tilde{\Gamma}^f_t,\end{aligned}$$

with respect to the semi-distance $\overleftarrow{\eta}(\gamma,\Lambda)=\min_{\lambda\in\Lambda}|\gamma-\lambda|$. The expansion process halts at the increment $d\tilde{\Gamma}^f_t$ satisfying $\max_{\gamma^f\in d\tilde{\Gamma}^f_t}\overleftarrow{\eta}\left(\gamma^f_t,\tilde{\Gamma}^f_t\right)<2/\rho$.

(ii) Vertex Selection: To minimize the *as is* primary set, next, the set of vertices $\hat{\Gamma}^f=\{\hat{\gamma}_k\in\tilde{\Gamma}^f\}$ satisfying the following conditions are selected:

$$\begin{aligned}\forall m,k:\quad \theta_{mk}-\theta_{nk} &< \pi, \qquad (6)\\ \hat{\gamma}_{(\cdot)}-\hat{\gamma}_k &= \left|\hat{\gamma}_{(\cdot)}-\hat{\gamma}_k\right|e^{j\left(\theta_{(\cdot)k}+\theta_k\right)},\\ &\quad \hat{\gamma}_{(\cdot)}\in\tilde{\Gamma}^f,\hat{\gamma}_k=|\hat{\gamma}_k|\,e^{j\theta_k}.\end{aligned}$$

(iii) *As-is* Primary Separation: Finally, the distribution of $\hat{\Gamma}$ is expanded along the following repulsive force:

$$d\hat{\gamma}_k = \sum_{\hat{\gamma}_j\in\hat{\Gamma}}(\hat{\gamma}_k-\hat{\gamma}_j)\,g_\alpha(\phi_k|\phi_j), \qquad (7)$$

within the possible coloring circle $|\hat{\gamma}_k|\leq 1$. As the results, the vertices $\{\hat{\gamma}_k\}$ are separated each other to yield a set of *as-is* primaries $\hat{\Pi}$ in Γ.

The effectiveness of the *as-is* primaries is illustrated in the main window of Fig. 5 where four instances of *as-is* primaries $\hat{\Pi}=\{\hat{\pi}_\text{r},\hat{\pi}_\text{g},\hat{\pi}_\text{gb},\hat{\pi}_\text{b}\}$ are estimated to implement the following *as-is* saliency filter:

$$\begin{aligned}\hat{\psi}_\omega &= \exp\left[-\hat{\mathcal{H}}_\omega\right], \qquad (8)\\ \hat{\mathcal{H}}_\omega &= -\sum_{\hat{\pi}_i\in\hat{\Pi}}p(\phi_\omega|\hat{\pi}_i)\log p(\phi_\omega|\hat{\pi}_i), \quad \gamma_\omega=e^{\text{RGB}}\phi_\omega,\end{aligned}$$

with respect to the primary selection probability given by

$$p(\phi_\omega|\hat{\pi}_i) = \frac{g_\alpha(\phi_\omega|\hat{\pi}_i)}{\sum_{\hat{\pi}_i\in\hat{\Pi}}g_\alpha(\phi_\omega|\hat{\pi}_i)}.$$

In this equation, $\hat{\pi}_i$ is restored via the following procedure:

$$\hat{\pi}_i = \tilde{\pi}_i+\overline{\pi}_i 1^{\text{RGB}}, \qquad 1^{\text{RGB}}=[1\quad 1\quad 1]^T, \qquad (9)$$

where $\tilde{\pi}_i=\frac{2}{3}\left(e^{\text{RGB}}\right)^T\hat{\gamma}_i$ and $\overline{\pi}_i$ designates a nominal brightness level given as the solution to the following

$$3\overline{\pi}_i^2+2\tilde{\pi}_i^T 1^{\text{RGB}}\cdot\overline{\pi}_i+|\tilde{\pi}_i|^2 = 1.$$

The implication of the chromatic complexity generator is demonstrated in the subwindows of Fig. 5; the distribution of the samples $\mathfrak{s}$ and the associated field $\varphi_\rho(\gamma|\mathfrak{s})$ are displayed in upper and lower subwindows, respectively. Through the comparison of Figs. 5 with 3, the *as-is* primary is effective to extend the focus of perception channel to landmark patterns wrapped by 'natural colors'.

4. MULTI-FRACTAL ARTICULATION

For the set of the chromatic information $\mathfrak{s}=\{\phi_i\}$ with size $\|\mathfrak{s}\|$, define the local diversity parameter by

$$\begin{aligned}\mathrm{R}_\mathfrak{s} &= \frac{1}{2\pi\alpha}\exp\left[-\frac{\sigma^2_{\phi\phi}}{2\alpha}\right],\\ \sigma^2_{\phi\phi} &= \frac{1}{\|\mathfrak{s}\|(\|\mathfrak{s}\|-1)}\times\sum_{\substack{1\leq i,j\leq\|\mathfrak{s}\|\\ i\neq j}}\left|\phi_i-\phi_j\right|^2.\end{aligned}$$

Then we can introduce a matching rule for selecting the pixels supervenient to an *as-is* primary:

$$g_\alpha(\omega|\hat{\pi}_i)>\mathrm{R}_\mathfrak{s} \;\Rightarrow\; \phi_\omega\overset{\mathrm{R}_\mathfrak{s}}{\sim}\hat{\pi}_i. \qquad (10)$$

Let $\mathfrak{O}$ be such a saliency pattern given by

$$\mathfrak{O} = \left\{\omega\in\Omega\;\middle|\;\phi_\omega\overset{\mathrm{R}_\mathfrak{s}}{\sim}\hat{\pi}_i\in\hat{\Pi}\right\}, \qquad (11)$$

and suppose that a fractal attractor Ξ is generated via successive applications of the contraction mappings $\mu_i:\Omega\mapsto\Omega$ of the following form:

$$\mu_i(\omega) = \frac{1}{2}[\omega+\omega^f_{\mu_i}], \qquad i=1,2,\ldots,\|\nu\|, \quad (12)$$

where $\omega^f_{\mu_i}$ is associated fixed point. Noticing that the fixed points should be located the Laplacian-Gaussian boundary $\partial^g\mathfrak{O}$, we can apply the successive expansion scheme (5) to the saliency pattern, as well. For articulating the set $\partial^g\mathfrak{O}$ with respect to the saliency patterns, a fixed point in the set $\tilde{\Omega}^f$ should be expanded via the following unification scheme:

$$\begin{aligned}\Omega^*_{t+1} &= \Omega^*_t\cup d\Omega^*_t, \qquad (13)\\ d\Omega^*_t &= \left\{\partial\omega^*\;\middle|\;\forall\partial\omega:\overleftarrow{\eta}\left(\partial\omega^*,\Omega^*_t\right)\leq\overleftarrow{\eta}\left(\partial\omega,\Omega^*_t\right)\right\},\end{aligned}$$

Figure 6: Fractal Articulation in Image Space
a saliency pattern associated with $\hat{\pi}_{\mathtt{r}}$-primary (Fig. 5) is articulated and allocated in the ground-object structure (Fig. 1).

where $\partial\omega^*, \partial\omega \in \Omega_T^f - \Omega_t^*$. To confine the attractor within a saliency pattern, the process (13) is interrupted at the increment $d\Omega_t^*$ satisfying the breakdown condition $\min_{\xi\in\Xi} g_\alpha(\xi|\hat{\pi}_i) < \mathtt{R}_\mathfrak{s}$.

Figure 6 shows the result of the fractal articulation process applied to the scene image (Fig. 1); in response to the selection of an *as-is* primary $\hat{\pi}_{\mathtt{r}} \in \hat{\Pi}$, the distribution of saliency colors $\mathfrak{O}$ is extracted and articulated via the successive procedure (13); resulted attractor is displayed in a subwindow localized within the perspective induced by the ground-object structure.

5. EXPERIMENTS AND DISCUSSIONS

The fractal articulation process is effective to separate a landmark object in complex background including distractive objects wrapped by similar colors as shown in Fig. 7. In this close-up view of the ground-object structure (a), 995 samples of unique chromatic information was collected to extract a saliency pattern $\mathfrak{O}$ with respect to a shifted and splitted version of *as-is* primary $\hat{\Pi} = \{\hat{\pi}_{\mathtt{r}}, \hat{\pi}_{\mathtt{rg}}, \hat{\pi}_{\mathtt{g}}, \hat{\pi}_{\mathtt{gb}}, \hat{\pi}_{\mathtt{b}}\}$ (b); the fixed points are successively articulated to specify a chromatically consistent subset $\hat{\Omega}^f$ at the first interruption (c); the fractal attractor controlled by the fixed points is indicated in the perspective of the scene (d). Figures 8 and 9 show the results of other experiments. In these experiments, saliency patterns are extracted as indicated in (a); associated fractal codes are visualized in the ground-object structure as illustrated in (b). These experimental results demonstrate that the fractal coding scheme is effective for the structurization of even ill-conditioned scenes where landmark objects are not observed as dominant patterns; a warning color of vehicles sometimes smaller image than attractive distractions as displayed in Fig. 8; the post is rather 'low-keyed' object in a night view as shown in Fig. 9. Even in such scene images, shifted version of the *as-is* primary is effectively estimated to generate $\hat{\psi}_\omega f_\omega^{\mathrm{RGB}}$-image.

The essential part of the $\hat{\psi}_\omega$-filtering is formalized by one dimensional successive manipulation (5)+ (6)+ (7) of finite features $\tilde{\Gamma}^f$. Noting that the features are extracted via parallel distributed system (4), we can implement the object detection scheme consisting of locally parallel processors.

The effectiveness of saliency-sensitive image analysis crucially dependent on the selection of the primary system. Through experimental studies using various types of scene

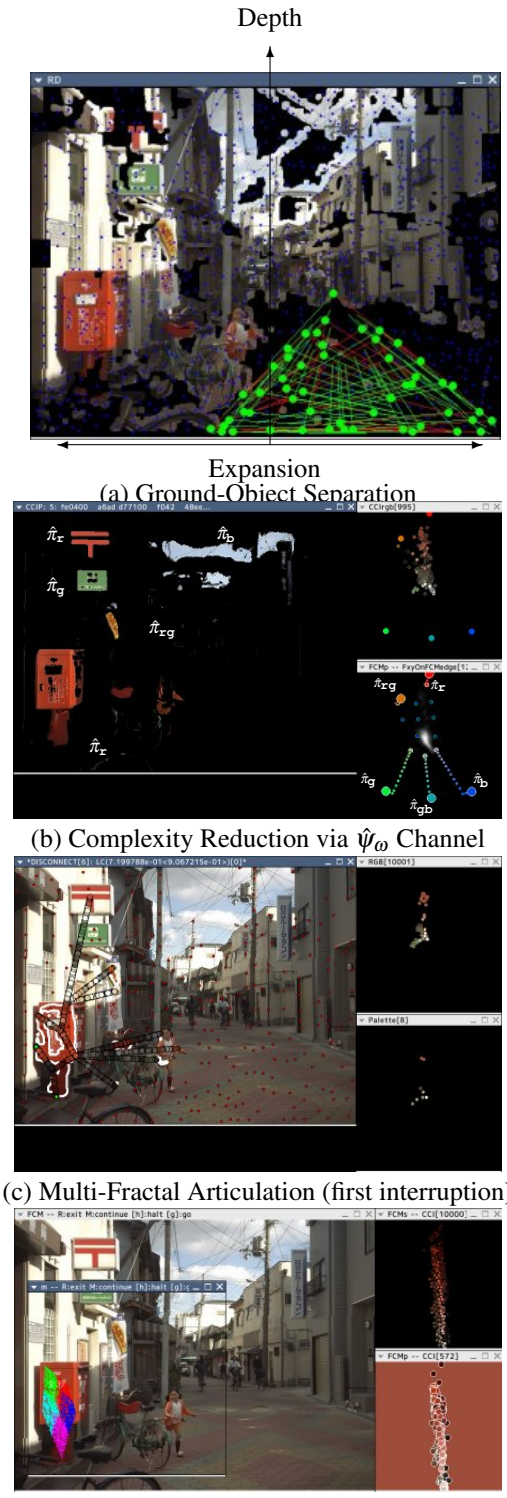


Figure 7: Fractal Coding of Scene Image
saliency patterns associated with as-is primary are articulated into a fractal attractors to allocate a landmark in the scene image.

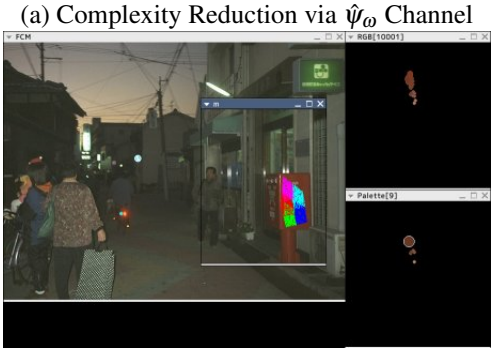
images, it has been demonstrated that the *as-is* primary is well instantiated to restore saliency colors missed and/or degenerated in conventional RGB-based image analysis. The

(a) Complexity Reduction via $\hat{\psi}_\omega$ Channel

(b) Contextual Visualization

Figure 8: Fractal Coding of Scene Image
saliency patterns associated with $\hat{\pi}_r$-primary (a) are articulated into a fractal attractors to allocate a landmark (b).

(a) Complexity Reduction via $\hat{\psi}_\omega$ Channel



(b) Contextual Visualization

Figure 9: Fractal Coding of Scene Image
saliency patterns associated with $\hat{\pi}_r$-primary (a) are articulated into a fractal attractors to allocate a landmark (b).

significance of the $\hat{\psi}_\omega$-filtering is summarized in Table 1 where the reduction of the unpredictability is evaluated in terms of relative complexity index $d\mathcal{S}_{(\cdot)} = \mathcal{S}_\emptyset - \mathcal{S}_{(\cdot)}$; $\mathcal{S}_\emptyset$ and $\mathcal{S}_G$ designate the Shannon's entropy with respect to the uniform distribution and the gray level distribution of f_ω, respectively; $\mathcal{S}_{\mathcal{H}}$ and $\mathcal{S}_{\hat{\mathcal{H}}}$ stand for the entropy with respect to the normalized versions of the distributions ψ_ω and $\hat{\psi}_\omega$, respectively. As shown in Table 1, the essential length of decision steps on the $\hat{\psi}_\omega f_\omega^{\mathrm{RGB}}$-image is reduced to $e^{-d\mathcal{S}_{\hat{\mathcal{H}}}} = 1/3 \sim 1/24$ of random search; except for distractive scene (Fig. 8), the computational complexity based on the *as-is* primary is no greater than 40% of the decision steps by using conventional trichromatic system.

Table 1: Complexity Reduction via $\hat{\psi}_\omega$-Filtering

scene	$d\mathcal{S}_G$	$d\mathcal{S}_{\mathcal{H}}$	$d\mathcal{S}_{\hat{\mathcal{H}}}$	$\|\hat{\Pi}\|$
Figs. 3, 5	0.131512	0.995172	2.032901	4
Fig. 7	0.199979	2.016679	2.884059	5
Fig. 8	0.103411	1.831957	2.059536	3
Fig. 9	0.148436	1.481229	3.189055	3
Fig. 2	0.196245	0.616455	1.266864	4

6. CONCLUDING REMARKS

A multi-fractal coding was applied to saliency based object detection in naturally complex scenes. By identifying the chromatic diversity with a fractal attractor in a color space, an *as-is* primary system is estimated to extract saliency patterns. Via multi-fractal coding of the saliency patterns, landmark objects are detected and allocated within the perspective of the ground-object structure. The predictability of the *as-is* primary in the variation of ambient light is left to future investigations.

REFERENCES

[1] J. J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin Company, Boston, Massachusetts, 1979.

[2] C. Hagelhall, T. Purcell, and T. R. P. Fractal dimension of landscape silhouette as a predictor for landscape preference. *Journal of Emvironmental Psychology*, 24:247–255, 2004.

[3] L. Itti, C. Kock, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-20(11):1254–1259, 1998.

[4] K. Kamejima. Anticipative coding and *In-Situ* adaptation of maneuvering affordance in a naturally complex scene. In V. A. Kulyukin, editor, *Advances in Human-Robot Interaction*, chapter 19, pages 307–324. In-Teh, Vukovar, Croatia, 2010.

[5] B. A. Olshausen, C. H. Anderson, and D. C. Van Essen. A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience*, 13(11):4700–4719, 1993.