International Journal of Innovative
Computing, Information and Control
Volume 8, Number 3(B), March 2012
 ISSN 1349-4198
pp. 2233–2248

MULTI-FRACTAL ARTICULATION OF ENVIRONMENTAL SALIENCY FOR CONTEXTUAL VISUALIZATION

Kohji Kamejima

Faculty of Information Science and Technology
Osaka Institute of Technology
1-79-1 Kitayama, Hirakata, Japan
kamejima@is.oit.ac.jp

Received March 2011; revised July 2011

Abstract. *A new framework is presented for saliency based object detection in naturally complex scenes. By identifying the ground-object structure through the scene images, in this framework, the random distribution of chromatic information is articulated into saliency patterns with respect to a scene specific primary system. Due to the intrinsic coherence of the scale and chromatic randomness, detected saliency patterns are well localized as object images within the perspective induced by the ground-object structure. By using the saliency pattern, the computational complexity of landmark detection can be significantly reduced.*
Keywords: Environmental saliency, Scale-chromatic complexity, Multi-fractal articulation

1. **Introductory Remarks.** Despite the infinite diversity of their appearance, natural scenes exhibit environment specific landmarks to be identified within individual intention of viewers. To control the focus on such a landmark object [18], perception processes should gather randomly distributed image features and apply 'feature integration' schemes [24]. In a conventional integration scheme, observed images are assumed to be structured in terms of 'visible' representations including key-points of well-organized objects and chunks of attractive colors; such visible saliency is computationally associated with the landmarks in complex scenes [7]. In many natural scenes, the visible saliency should support a multitude of viewer specific decision makings. However, the maintenance of the consistency is still an open problem in the representation of the visible saliency as a transversal support. In addition, practical scene images exhibit superfluous saliency patterns for individual decision makers. In such a situation, deterministic integration of image features easily incurs combinatorial explosion.

In many practical situations, the perspective from each vehicle is restricted by boundary objects surrounding the roadway area. Despite such physical-geometric restrictions, recent advancements of space technology combined with large scale information networks provide a logical-computational basis for over the horizon of visible scenes. For instance, current global positioning system (GPS) yields effective information for dynamic localization of vehicles along roadway areas [28, 30]. The behaviors of such vehicles are matched with geometric representations of local terrains for planning [3], regulating [32] and operating [27, 33] vehicle control processes equipped with self-reliant intelligence. By networking GPS and vehicle control systems with the earth observation systems, additionally, we can detect a roadway pattern beyond the physical-geometric perspective [15]. The random distribution of scale feature is extracted in a scene image to match with a generic roadway model for sampling a 'palette'; the palette is transferred to a cut of a satellite image for

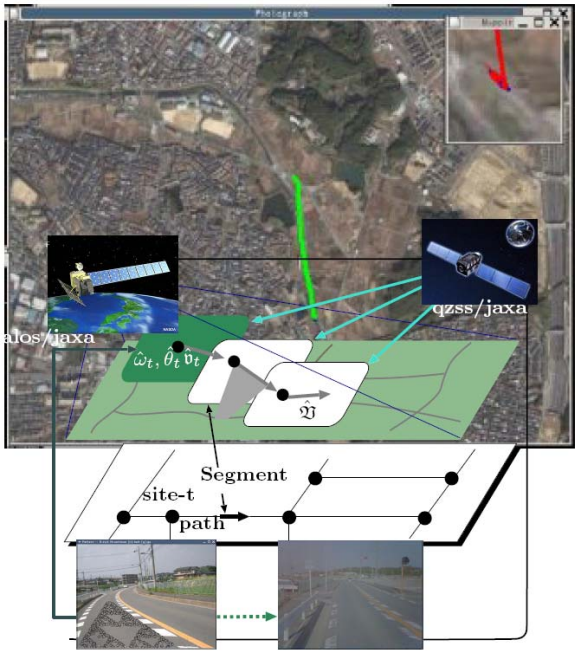


FIGURE 1. Anticipative road following system

the detection and extension of a roadway pattern as illustrated in Figure 1. Let $\left(\hat{\omega}_t, \hat{\theta}_t\right)$ be the estimate of the current position and direction in the satellite image with local roadway segment $\hat{\mathfrak{v}}_t$. Suppose that the local segment is extended to generate a chain $\hat{\mathfrak{V}} = \{\, \hat{\mathfrak{v}}_t, t = 1, 2, \dots \}$ to a possible destination in the bird's eye view. Through such an 'anticipative' road following, the geometric complexity of the local terrain is reduced to a graph on which we can constrain the viewpoint of future scenes *prior to* physical access. By using the graph structure, we can extend the scope of geometric planning [3, 17] and autonomous vehicle guidance [8, 31] to the over-the-horizon maneuvering processes. However, due to essentially unpredictable distribution of temporal and/or moving objects, it is not practical to utilize the satellite image as the reference of the vehicle control for a subsequent maneuvering process. This implies that the decision making by each advanced vehicles must be supported by the information to be finally gathered via on-vehicle vision systems. As a result, the scope of vehicle control systems is still restricted within a physical-geometric perspective.

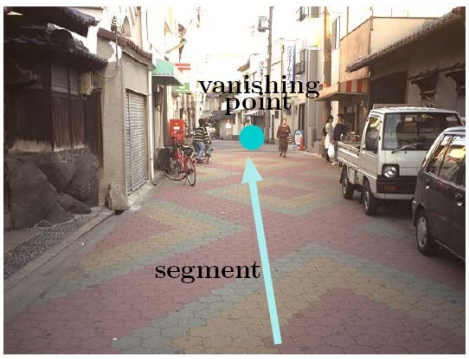


FIGURE 2. Scene to be analyzed

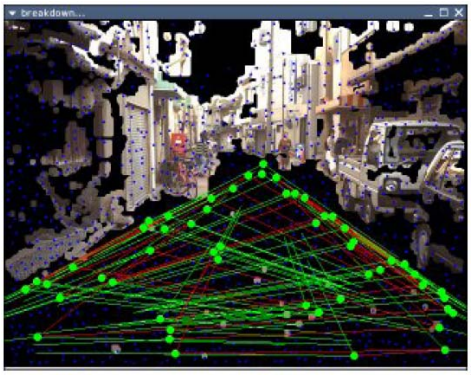


FIGURE 3. Fractal model of roadway area

In this paper, we have introduced a new framework for identifying vehicle specific scenes based on a transferable representation of the roadway pattern. We assumed that the 'future trajectory' along the graph structure is downloaded to the on-vehicle vision as an *a priori* description of the scene to be observed. Following empirical knowledge of ecological optics [5] and neural-psychology [4, 25] combined with recent advancements in emotional – as well as computational-perception [6, 11], vision systems should organize the generic structure of maneuverable scenes in terms of fractal codes specifying the expansion of horizontal planes and the aggregation of boundary objects. Noting this, the randomness of the gray level distribution is extracted as a version of 'latent images' to estimate the fractal code spanning a connected open space [15]; the downloaded segment $\hat{\mathfrak{v}}_t$ is mapped into the scene image as shown in Figure 2; and, the random distribution of scale information is analyzed to design a fractal model confined by boundary objects as illustrated in Figure 3. In this figure, the expansion of the fractal model is indicated by a closed graph spanning a connected open space. Simultaneously, the aggregation of the boundary objects is visualized as the distribution of saliency colors. In order to operate on-vehicle vision in cooperation with human's inherent perception, thus, the saliency distribution should be articulated into a system of fractal codes spanning object images within the ground-object structure.

2. **Existence of Environmental Saliency.**

Through recent investigations in genetic physiology, it has been pointed out that inherent vision is equipped with a not-yet-explicated scheme for understanding the real world thronged with uproarious luminescence by friendly or undesirable neighbors [29]. Under the assumption that the environment is organized as the collaboration of intentional and/or contingent participants, we introduce a preestablished restriction of image features arising in 'naturally complex' scenes. To maintain the multitude of mutual coexistence, the participants should substantiate warning sign and/or informative design to be accepted as 'landmark objects' [19]. Knowing that objects which really exist should be degenerated through the iteration of physical deformation processes, simultaneously, the participants have competitively developed sophisticated sensing devices including the eyes, in particular. As a part of the actual world, thus, resulted environment of the co-evolution cycles should present all participants with a 'readable image' of the saliency sign/design. To control the focus of vision systems in accordance with the individual decision processes, it is pertinent to transform the image features into context free representations of the saliency sign/design called 'environmental saliency' [16]. Noting the apparent diversity of the real world, the environmental saliency should be grounded through stochastic aspects underlying various types of scene images [12]: the fluctuation of the local scale shift and the uncertainty of subtle chromatic scattering, e.g.

As intentional participants of the co-evolution process without the 'intelligent designer', it should be an essential capability for the humans to control the focus of the inherent vision within the surroundings [26] on the premise that imminent decision making should be evoked by 'light speed' transformation of unstructured ambient light [5]. Following the anatomy of early vision systems, the transformation is captured by two neuronal processes sensitive to brightness distribution and subtle to spectrum shift, respectively. It has been revealed that the multi-scale Gaussian filter is implemented by inherent parallel computation mechanism [23, 9] to extract 2.5 dimensional cue to the object detection [21, 22]. This implies that even weak light is sufficient to generate the following distributed

feature in early scene analysis [11]:

$$\hat{\sigma}_\omega = \sqrt{\frac{2f_\omega}{|\Delta f_\omega|}}, \tag{1}$$

where f_ω denotes the brightness of the incoming light at the pixel ω in an image plane Ω. The index $\hat{\sigma}_\omega$ provides a pixel wise evaluation of the upper bound of local scale information.

Such a 'fast' scale information has been exploited as a cue to early structural analysis including ground-object separation of complex scene imagery [14]. By matching the scale information $\hat{\sigma}_\omega$ with a generic perspective model, we have a probability distribution of the roadway area as shown in Figure 3. Simultaneously, the same optic array is accepted by a more sophisticated system where spectral diversity of incoming light is factorized within a system of primary colors [2]. Despite the loss of the depth information, saliency patterns extracted as 'matted' images [20] should be associated with 3D objects under universal preference to a class of fractal patterns.

3. **Locally Gaussian Palette.** As the cue to the multi-fractal coding, in what follows, another version of latent images, i.e., random distribution of chromatic complexity, is extracted as the observables of not-yet-identified landmark objects. Let $\mathtt{R}$, $\mathtt{G}$, $\mathtt{B}$ be three primaries, and suppose that the incoming light is identified with a linear combination of the primaries by humans. This implies that the information conveyed by the spectral distribution be described within the nonnegative subset of 3D Euclid space R^+_3 as follows:

$$f^{\mathtt{RGB}}_\omega = \begin{bmatrix} f^{\mathtt{R}}_\omega & f^{\mathtt{G}}_\omega & f^{\mathtt{B}}_\omega \end{bmatrix}^T,$$

where $f^{(\cdot)}_\omega$ designates subjective weight of the primary $(\cdot)$. Define $\phi_\omega = f^{\mathtt{RGB}}_\omega / |f^{\mathtt{RGB}}_\omega|$. By identifying the totality of the chromatic information ϕ_ω with the positive part of the unit sphere ∂^3_+, we can induce the following measure:

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

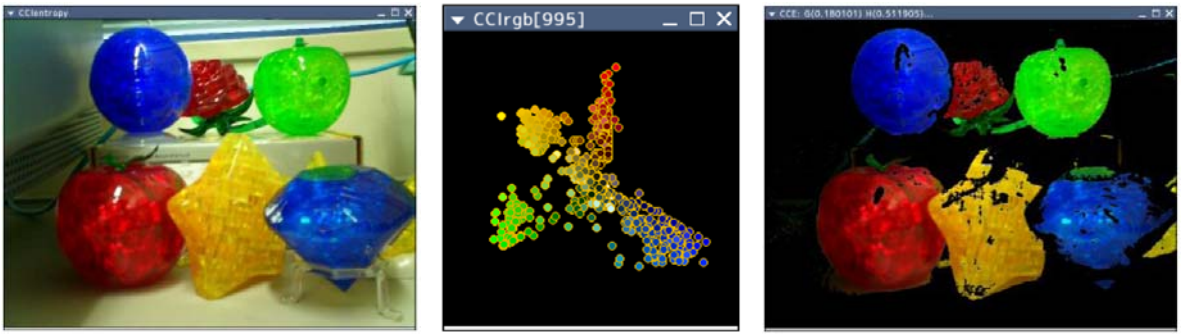
Following experimental studies using various types of roadway scenes, the sensivity factor α should be adjusted to $1/10 \sim 1/100$ [13]. For sufficiently small chromatic variation $|\phi - \phi_\omega|$, the measure $g_\alpha(\phi|\phi_\omega)$ approximates the Gaussian distribution on local tangential space at ϕ_ω. Noticing the following constraint,

$$|\phi_\omega|^2 = \sum_{\mathtt{RGB}} \left(\frac{f^{(\cdot)}_\omega}{|f^{\mathtt{RGB}}_\omega|}\right)^2 = 1,$$

by definition, we have the following index for evaluating coloring saliency

$$\begin{aligned} \psi_\omega &= e^{-\mathcal{H}_\omega}, \\ \mathcal{H}_\omega &= -2\sum_{\mathtt{RGB}} \left(\frac{f^{(\cdot)}_\omega}{|f^{\mathtt{RGB}}_\omega|}\right)^2 \log\left(\frac{f^{(\cdot)}_\omega}{|f^{\mathtt{RGB}}_\omega|}\right). \end{aligned} \tag{3}$$

In this indexing, the substantial process in the retina system is considered to generate the 'square root of capturing probability' $\sim \phi_\omega$ to yield the random distribution of Shannon's entropy $\mathcal{H}_\omega$; the complexity measure is transformed to a saliency probability ψ_ω via vitals specific 'neg-entropy' generation. By using the saliency indexing (12), we can detect landmark objects in a well-structured scene as illustrated in Figure 4, where the chromatic complexity (= simplicity, in this case) of the 'block world' (a) is represented in



(a) A block world (b) RGB color space (c) Filtered image

FIGURE 4. Complexity reduction via ψ_ω channel (3d puzzle)

FIGURE 5. Complexity reduction via ψ_ω channel (shopping street)

a conventional color space (b) for 'matting' the saliency patterns in a noisy background (c).

The saliency indexing can be applied to a naturally complex scene to visualize the distribution $\psi_\omega f_\omega^{\mathtt{RGB}}$ as shown in Figure 5, where the complexity of the scene indicated in Figure 2 is reduced via the ψ_ω-filtering to visualize a partial distribution of landmark objects. As demonstrated in Figures 2 and 5, the saliency probability is sensitive to the objects marked by preassigned primaries. To correct such pigmentation-level bias, we need adaptation of the primary system to the entire the chromatic complexity arising in the observed scene.

4. **Chromatic Complexity Generator.** Figure 4 implies the existence of a chromatic complexity generator. The diversity of the chromatic information is expanded towards the set of the primaries. Noticing this, consider an inverse problem: how to select a set of fixed points to regenerate the chromatic diversity as a fractal attractor in the color space. For such a global analysis of chromatic diversity, let the chromatic complexity index ϕ_ω be identified with the following planar representation of the color space:

$$\Gamma \ni \gamma = e^{\mathtt{RGB}}\phi_\omega, \qquad e^{\mathtt{RGB}} = \left[e^{\mathtt{R}} \quad e^{\mathtt{G}} \quad e^{\mathtt{B}}\right], \qquad e^{(\cdot)} = \left[\cos\theta_{(\cdot)} \quad \sin\theta_{(\cdot)}\right]^T, \tag{4}$$

with *a priori* orientation $\theta_{\mathtt{R}} = \pi/2$, $\theta_{\mathtt{G(B)}} = \theta_{\mathtt{R}} + (-)2\pi/3$. The diversity of the incoming light $f_\omega^{\mathtt{RGB}}$ is represented in the color space Γ through the linear transform (4). The representation in Γ can be exploited to restore the chromatic information through the following

procedure:

$$\hat{\phi}_\omega = \tilde{\phi}_\gamma + \overline{\phi}_\omega 1^{\mathtt{RGB}}, \qquad 1^{\mathtt{RGB}} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T, \tag{5}$$

where $\tilde{\phi}_\gamma = \dfrac{2}{3}\left(e^{\mathtt{RGB}}\right)^T \gamma$. In (5), $\overline{\phi}_\omega$ designates a nominal brightness level given as the solution to the following equation

$$3\overline{\phi}_\omega^2 + 2\tilde{\phi}_\gamma^T 1^{\mathtt{RGB}} \cdot \overline{\phi}_\omega + \left|\tilde{\phi}_\gamma\right|^2 = 1.$$

By introducing the representation in Γ, we can enhance the spectrum shift independent of the variation of the brightness level in the observed imagery.

Suppose that samples of the chromatic information $f_{\omega_1}^{\mathtt{RGB}}, f_{\omega_2}^{\mathtt{RGB}}, \ldots, f_{\omega_n}^{\mathtt{RGB}}$, are collected in a scene image to generate a palette $\mathfrak{s} = \left\{ \phi\left(f_{\omega_i}^{\mathtt{RGB}}\right), \omega_i \in \Omega \right\}$ of the size $\|\mathfrak{s}\| = n$. Let $\chi_{\mathfrak{s}}$ be the aggregation of Dirac's delta measure distributed on the set $\Gamma_{\mathfrak{s}} = \left\{ \gamma\left(\phi_\omega\right) \;\middle|\; \phi_\omega \in \mathfrak{s} \right\}$ and consider the field information $\varphi_\rho\left(\gamma|\mathfrak{s}\right)$ generated via the following dynamics on Γ:

$$\frac{\partial}{\partial t}\varphi_\rho\left(\gamma|\mathfrak{s}\right) = \frac{1}{2}\Delta\varphi_\rho\left(\gamma|\mathfrak{s}\right) + \rho\left[\chi_{\mathfrak{s}} - \varphi_\rho\left(\gamma|\mathfrak{s}\right)\right], \tag{6}$$

where $\chi_{\mathfrak{s}}$ denotes the aggregation of Dirac's delta measure distributed on the set $\Gamma_{\mathfrak{s}} = \left\{ \gamma\left(\phi_\omega\right) \;\middle|\; \phi_\omega \in \mathfrak{s} \right\}$. Assume that the distribution $\Gamma_{\mathfrak{s}}$ is identified with a degenerated version of $\Xi_{\mathfrak{s}}$; a fractal attractor satisfying the following constraint:

$$\Xi_{\mathfrak{s}} \;=\; \bigcup_{\hat{\pi}_i \in \hat{\Pi}} \mu_{\hat{\pi}_i}\left(\Xi_{\mathfrak{s}}\right), \tag{7}$$

with respect to an *as-is* set of scene specific primary $\hat{\Pi} = \{ \hat{\pi}_i \}$; $\mu_{\hat{\pi}_i}$ denotes a contraction mapping $\Gamma \mapsto \Gamma$ with a fixed point $\hat{\pi}_i$. Equation (7) implies that the intrinsic diversity of the chromatic information is re-generated via the iterated function system [1] essentially governed by the allocation of the fixed points $\hat{\Pi}$. By adjusting the system (6) to the complexity of the iterated function system, i.e., $\rho = \log_2 \|\hat{\Pi}\|$, we can evaluate the probability for capturing the not-yet-identified attractor $\Xi_{\mathfrak{s}}$ in terms of the solution $\varphi_\rho\left(\omega|\mathfrak{s}\right)$ [10]. Noticing that the fixed point must be allocated on the Laplacian-Gaussian boundary $\partial_{\mathfrak{s}}^g$ specified by the capturing probability $\varphi_\rho\left(\gamma|\mathfrak{s}\right)$, we can identify an *as-is* primary $\hat{\Pi}$ based on the distribution $\chi_{\mathfrak{s}}$. The identification process is divided into the following three steps. First, a possible fixed point $\tilde{\gamma}_0^f$ is located on the Laplacian-Gaussian boundary $\partial_{\mathfrak{s}}^g$ and expanded via the following successive scheme:

$$\tilde{\Gamma}_{t+1}^f = \tilde{\Gamma}_t^f \cup d\tilde{\Gamma}_t^f, \qquad \tilde{\Gamma}_0^f = \left\{ \tilde{\gamma}_0^f \right\}, \tag{8}$$

where the increment is selected as follows:

$$\begin{aligned} d\tilde{\Gamma}_t^f \;=\; & \left\{ \tilde{\partial}\gamma^f \;\middle|\; \forall \partial\tilde{\gamma} : \overleftarrow{\eta}\left(\partial\tilde{\gamma}^f, \tilde{\Gamma}_t^f\right) \geq \overleftarrow{\eta}\left(\partial\tilde{\gamma}, \tilde{\Gamma}_t^f\right) \right\}, \\ & \tilde{\partial}\gamma^f, \partial\tilde{\gamma} \in \partial_{\mathfrak{s}}^g - \tilde{\Gamma}_t^f, \end{aligned}$$

with respect to $\overleftarrow{\eta}\left(\gamma, \Lambda\right) = \min_{\lambda \in \Lambda} |\gamma - \lambda|$. This expansion process mutually separates the fixed points in the boundary set $\partial_{\mathfrak{s}}^g$ and halts at T when the increment $d\tilde{\Gamma}_T^f$ satisfies the following sub-scale condition:

$$\max_{\gamma \in d\tilde{\Gamma}_T^f} \eta\left(\gamma, \tilde{\Gamma}_T^f\right) < \sqrt{1/\rho}.$$

Next, a subset $\hat{\Gamma} = \{\hat{\gamma}_k\} \subset \tilde{\Gamma}_T^f$ satisfying the following conditions is selected as an estimate of the vertices

$$\begin{aligned} \forall m,k: \quad & \theta_{mk} - \theta_{nk} < \pi, \\ & \hat{\gamma}_{(\cdot)} - \hat{\gamma}_k = \left|\hat{\gamma}_{(\cdot)} - \hat{\gamma}_k\right| e^{j\left(\theta_{(\cdot)k}+\theta_k\right)}, \\ & \hat{\gamma}_{(\cdot)} \in \tilde{\Gamma}_T^f, \quad \hat{\gamma}_k = |\hat{\gamma}_k| \, e^{j\theta_k}. \end{aligned} \tag{9}$$

Selected vertex $\hat{\gamma}_k \in \hat{\Gamma}$ is considered to be a degenerated version of the *as-is* primary $\hat{\pi}_k$ resulted from the following process:

$$\hat{\pi}_k = \check{\gamma}_0^k \to \check{\gamma}_1^k \to \check{\gamma}_2^k \to \cdots \to \hat{\gamma}_k, \tag{10}$$

where $\check{\gamma}_t^k$ is t-step degeneration of $\hat{\pi}_k$.

Finally, the distribution $\hat{\Gamma}$ is expanded to simulate the anti-degeneration process. In the simulated process, the *as-is* primary $\hat{\pi}_k = \check{\gamma}_0^k$ is successively restored by using the mutually repulsive force, i.e.,

$$\begin{aligned} \check{\gamma}_t^k &= \check{\gamma}_{t+1}^k + d\check{\gamma}_k, \\ d\check{\gamma}_k &= \kappa \sum_{\tilde{\gamma}_j \in \tilde{\Gamma}} \left(\check{\gamma}_k - \tilde{\gamma}_j\right) g_\alpha \left(\check{\phi}_k | \tilde{\phi}_j\right), \end{aligned} \tag{11}$$

with positive control parameter κ. In Equation (11), $\check{\phi}_k\left(\tilde{\phi}_j\right)$ denotes the restored chromatic information of $\check{\gamma}_k\left(\tilde{\gamma}_j\right)$, i.e.,

$$\begin{aligned} \check{\phi}_t^k &= \frac{2}{3}\left(e^{\mathtt{RGB}}\right)^T \check{\gamma}_t^k + \overline{\phi}_t^k 1^{\mathtt{RGB}}, \\ & 3\left(\overline{\phi}_t^k\right)^2 + \frac{4}{3}\left(\check{\gamma}_t^k\right)^T e^{\mathtt{RGB}} 1^{\mathtt{RGB}} \cdot \overline{\phi}_t^k + \left|\frac{2}{3}\left(e^{\mathtt{RGB}}\right)^T \check{\gamma}_t^k\right|^2 = 1, \end{aligned}$$

$$\begin{aligned} \tilde{\phi}_j &= \frac{2}{3}\left(e^{\mathtt{RGB}}\right)^T \tilde{\gamma}_j + \overline{\phi}_j 1^{\mathtt{RGB}}, \\ & 3\overline{\phi}_j^2 + \frac{4}{3}\tilde{\gamma}_j^T e^{\mathtt{RGB}} 1^{\mathtt{RGB}} \cdot \overline{\phi}_j + \left|\frac{2}{3}\left(e^{\mathtt{RGB}}\right)^T \tilde{\gamma}_j\right|^2 = 1, \end{aligned}$$

respectively. The simulated anti-degenerator (11) expands the distribution $\check{\Gamma}$ within the possible coloring circle $\left|\check{\gamma}_t^k\right| \leq 1$.

The scheme (8) combined with (9) yields a set of vertex points $\hat{\Gamma}$ to be associated with contraction mappings for regenerating the distribution $\Gamma_{\mathfrak{s}}$ as a degenerated version of the attractor $\Xi_{\mathfrak{s}}$. The combination with the dynamics (11) yields a generator of the fractal dynamics controlled by the *as-is* primary $\hat{\pi}_i = \check{\gamma}_0^i$ given by $\hat{\pi}_i = \tilde{\pi}_i + \overline{\pi}_i 1^{\mathtt{RGB}}$ where

$$\begin{aligned} \tilde{\pi}_i &= \frac{2}{3}\left(e^{\mathtt{RGB}}\right)^T \check{\gamma}_0^i, \\ & 3\overline{\pi}_i^2 + 2\tilde{\pi}_i^T 1^{\mathtt{RGB}} \cdot \overline{\pi}_i + |\tilde{\pi}_i|^2 = 1. \end{aligned}$$

By using the *as-is* primary, we can extend the saliency indexing (12) to naturally complex scenes. To this end, let the chromatic diversity within the scope of the degenerated primary $\hat{\Gamma}$ be indexed by

$$Q\left(\omega|\hat{\Gamma}\right) = \sum_{\hat{\gamma}_i \in \hat{\Gamma}} g_\alpha\left(\phi_\omega|\hat{\gamma}_i\right),$$

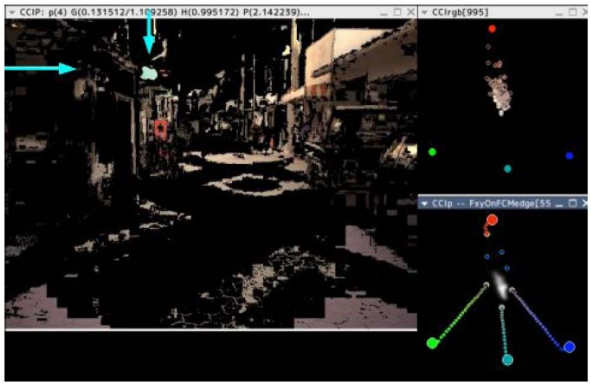


FIGURE 6. Complexity reduction via $\hat{\psi}_\omega$ channel (shopping street)

and define

$$\hat{\mathcal{H}}_\omega = -\sum_{\hat{\pi}_i \in \hat{\Pi}} g_\alpha\left(\omega|\hat{\pi}_i, \hat{\Pi}\right) \log g_\alpha\left(\omega|\hat{\pi}_i, \hat{\Pi}\right),$$

where $g_\alpha\left(\omega|\hat{\pi}_i, \hat{\Pi}\right)$ stands for the probabilistic complexity in the selection of *as-is* primary, i.e.,

$$\begin{aligned} g_\alpha\left(\omega|\hat{\pi}_i, \hat{\Pi}\right) &= \frac{1}{Q\left(\omega|\hat{\Pi}\right)} g_\alpha\left(\phi_\omega|\hat{\pi}_i\right), \\ Q\left(\omega|\hat{\Pi}\right) &= \sum_{\hat{\pi}_i \in \hat{\Pi}} g_\alpha\left(\phi_\omega|\hat{\pi}_i\right). \end{aligned}$$

Then, we have the following evaluation

$$\begin{aligned} \hat{\psi}_\omega &= e^{-\hat{\mathcal{H}}_\omega} Q\left(\omega|\hat{\Gamma}\right) \\ &= \frac{Q\left(\omega|\hat{\Gamma}\right)}{Q\left(\omega|\hat{\Pi}\right)} \exp\left[\frac{1}{Q\left(\omega|\hat{\Pi}\right)} \sum_{\hat{\pi}_i \in \hat{\Pi}} g_\alpha\left(\phi_\omega|\hat{\pi}_i\right) \log g_\alpha\left(\phi_\omega|\hat{\pi}_i\right)\right]. \end{aligned} \tag{12}$$

The chromatic complexity generator was applied to the scene image (Figure 2) to demonstrate the effectiveness of the *as-is* primaries as illustrated in Figure 6. The distribution of the samples $\mathfrak{s}$ and resulted estimates $\hat{\Pi}$ are indicated in upper and lower subwindows, respectively. As indicated in the lower subwindow, the *as-is* primary is split and adapted to the scene image. By designing $\hat{\psi}_\omega$-filtering based on the *as-is* primary $\hat{\Pi}$ and applying to $f_\omega^{\mathtt{RGB}}$-distribution in entire scene image, we have an enhancement of the saliency distribution. In the main window, the landmark objects are detected based on the *as-is* saliency index $\hat{\psi}_\omega$; a sign pattern marked by $\rightarrow$ and $\downarrow$ is enhanced as well as two types of landmarks of post office extracted in Figure 5. Through the comparison of Figures 5 and 6, it has been demonstrated that the *as-is* primaries is effective to extend the focus of perception channel to considerably degenerated saliency colors.

5. **Multi-Fractal Articulation.** The $\hat{\psi}_\omega$-filter can be exploited to classify the saliency distribution into a class of patterns $\{\mathfrak{O}_{\hat{\pi}_i}\}$ given by

$$\mathfrak{O}_{\hat{\pi}_i} = \left\{ \omega \in \Omega \;\middle|\; g_\alpha\left(\phi_\omega|\hat{\pi}_i\right) \geq g_\alpha\left(\phi_\omega|\hat{\pi}_j\right) \right\},$$

where $\hat{\pi}_i, \hat{\pi}_j \in \hat{\Pi}$. To detect landmark objects under the inherent preference to fractal attractors, let the saliency pattern $\mathfrak{O}_{\hat{\pi}_i}$ be articulated into self-similarity patterns satisfying the following constraint:

$$\Xi_{\hat{\pi}_i} = \bigcup_{\mu_i \in \nu} \mu_i \left(\Xi_{\hat{\pi}_i} \right), \tag{13a}$$

where ν denotes a set of contraction mappings $\mu_i : \Omega \mapsto \Omega$ of the following form:

$$\mu_i^{\Omega}(\omega) = \frac{1}{2} \left[\omega + \omega_{\mu_i}^f \right], \quad i = 1, 2, \ldots, \|\nu\|, \tag{13b}$$

with the fixed point $\omega_{\mu_i}^f$. Thus, the self-similarity process (13) is designable via the allocation of the fixed point $\omega_{\mu_i}^f$. The self-similarity mechanism (13) is visible on the scene image. By identifying the distribution $\hat{\psi}_\omega$ with the invariant measure of the self-similarity process, we can generate the probability distribution for capturing the attractor $\Xi_{\hat{\pi}_i}$ as the steady state of the following distributed parameter system [10]:

$$\frac{\partial}{\partial t} \varphi_b \left(\omega | \hat{\Pi} \right) = \frac{1}{2} \Delta \varphi_b \left(\omega | \hat{\Pi} \right) + \rho_b \left[\hat{\psi}_\omega - \varphi_b \left(\omega | \hat{\Pi} \right) \right]. \tag{14}$$

Following the empirical knowledge on the emotional perception [6], the complexity factor ρ_b should be adjusted to 1.3.

Equation (13) means that the self-similarity process is essentially designable via the allocation of the fixed point $\omega_{\mu_i}^f$. Noticing that the fixed points should be located the Laplacian-Gaussian boundary $\partial^g \mathfrak{O}_{\hat{\pi}_i}$ with respect to the capturing probability $\varphi_b \left(\omega | \hat{\Pi} \right)$, again, we can apply the following scheme for successive allocation of the fixed points to the saliency pattern, as well:

$$\begin{aligned} \tilde{\Omega}_{t+1}^f &= \tilde{\Omega}_t^f \cup d\tilde{\Omega}_t^f, \\ d\tilde{\Omega}_t^f &= \left\{ \tilde{\partial}\omega^* \mid \forall \tilde{\partial}\omega : \overleftarrow{\eta} \left(\tilde{\partial}\omega^*, \tilde{\Omega}_t^f \right) \geq \overleftarrow{\eta} \left(\tilde{\partial}\omega, \tilde{\Omega}_t^f \right) \right\}, \\ & \quad \tilde{\partial}\omega^*, \tilde{\partial}\omega \in \partial^g \mathfrak{O}_{\hat{\pi}_i} - \tilde{\Omega}_t^f. \end{aligned} \tag{15}$$

The expansion process (15) mutually separates the estimate of the fixed points $\tilde{\Omega}_t^f$ in the fixed set $\partial^g \mathfrak{O}_{\hat{\pi}_i}$. The allocation step halts at the increment $d\Omega_T^*$ satisfying

$$\max_{\omega \in d\Omega_T^*} \eta \left(\omega, \Omega_T^* \right) < \sqrt{1/\rho_b}, \tag{16}$$

with respect to the final set Ω_T^*.

Many physical entities can be surrounded by convex contours in the scene images. This implies that the final set Ω_T^* should be articulated according to a convexity criterion. To this end, we invoke the following nondeterministic algorithm:

$$\begin{aligned} \Omega_{t+1}^f &= \Omega_t^f \cup d\Omega_t^f, \\ d\Omega_t^f &= \left\{ \partial\omega^f \mid \overleftarrow{\eta} \left(\partial\omega^f, \Omega_t^f \right) \leq \overleftarrow{\eta} \left(\partial\omega, \Omega_t^f \right) \right\}, \\ & \quad \partial\omega^f, \partial\omega \in \Omega_T^* - \Omega_t^f. \end{aligned} \tag{17}$$

The successive process (17) expands the initial set $\Omega_0^f = \left\{ \omega_0^f \right\}$, $\omega_0^f \in \partial^g \mathfrak{O}_{\hat{\pi}_i}$ towards a convex set within the final expansion Ω_T^*. To maintain an on-going articulation Ω_t^f within

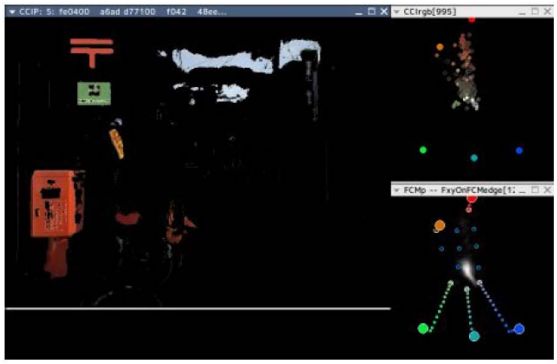


FIGURE 7. Complexity reduction via $\hat{\psi}_{\omega}$ channel (post office)

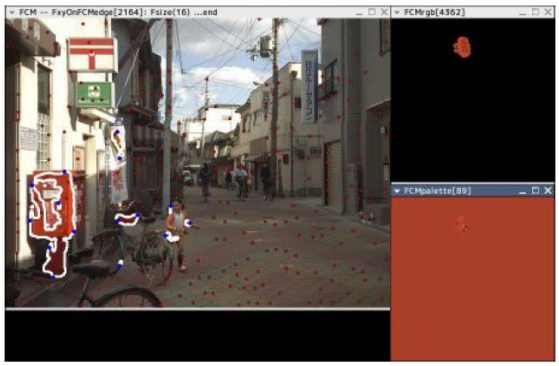


FIGURE 8. Fixed point allocation on $\partial^{g}\mathfrak{O}_{\hat{\pi}_i}$

a convex pattern, the process (17) is interrupted by the following breakdown criterion:

$$\min_{\xi \in d\Xi_t} g_{\alpha}\left(\phi_{\xi}|\mathfrak{s}\right) < \min_{\xi \in \Xi_t} g_{\alpha}\left(\phi_{\xi}|\mathfrak{s}\right), \tag{18}$$
$$g_{\alpha}\left((\cdot)|\mathfrak{s}\right) = \max_{f^{\mathtt{RGB}} \in \mathfrak{s}} g_{\alpha}\left((\cdot)|\phi\left(f^{\mathtt{RGB}}\right)\right),$$

where Ξ_t and $d\Xi_t$ denote fractal attractors associated with on-going and testing fixed points, i.e., Ω_t^f and

$$d\Omega_t^f \cup \left\{ \partial\omega^f \;\middle|\; \forall \partial\omega : \overleftarrow{\eta}\left(\partial\omega^f, d\Omega_t^f\right) \leq \overleftarrow{\eta}\left(\partial\omega, d\Omega_t^f\right) \right\},$$
$$\partial\omega^f, \partial\omega \in \Omega_t^f,$$

respectively. The articulation process is finally halted by the singularity condition: $\|\Omega_T^* - \Omega_t^f\| < 2$.

The multi-fractal articulation scheme was applied to the $\hat{\psi}_{\omega}$-filtering image as illustrated in Figure 7, where landmarks of a post office are emphasized in a close-up view of the scene image (Figure 2). By selecting an *as-is* primary of a dark-orange color, a saliency pattern $\mathfrak{O}_{\hat{\pi}_i}$ is extracted as indicated in Figure 8; white contours indicate the distribution of the Laplacian-Gaussian boundary $\partial^{g}\mathfrak{O}_{\hat{\pi}_i}$. On this boundary, the set of the fixed points $\tilde{\Omega}_T^*$ is estimated via the successive scheme (15) to articulate the saliency pattern as displayed in Figure 9. The fixed points in $\tilde{\Omega}_T^*$ are successively linked via the unification algorithm (17) towards convex patterns. The unification process is interrupted by the criterion (18) as indicated in Figure 9(a). At the first interruption, in this case, a fractal code ν is designed as an articulation of the saliency pattern $\mathfrak{O}_{\hat{\pi}_i}$; as indicated in Figure 9(b), the associated fractal attractor is generated for sampling the chromatic information on the

(a) First interruption (b) Fractal code

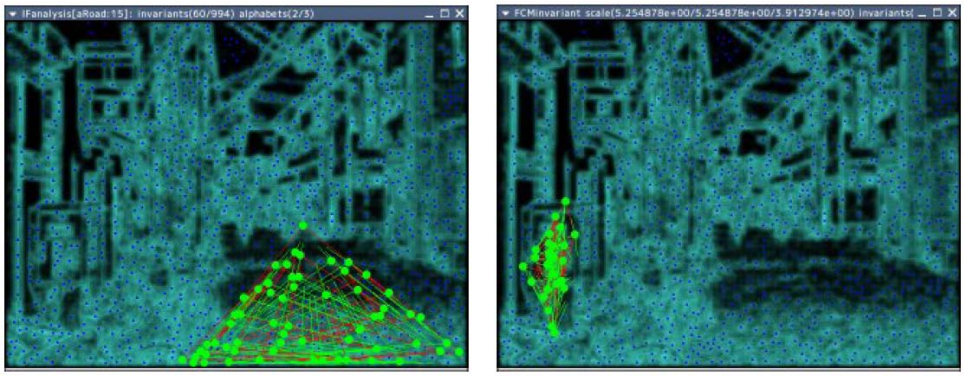
FIGURE 9. Articulation of boundary object

FIGURE 10. *A priori* structure of scene image (post office)

saliency pattern. As a result, we have a fractal model to be associated with a landmark of a post office. As demonstrated in Figure 9, we can exploit the breakdown condition (18) with respect to the *as-is* primary to articulate the saliency distribution into fractal attractors. The identification of the distribution with the fractal attractor implies that the connectedness of the articulated pattern is explicitly verified by the self-similarity (13).

The allocation algorithm (15) was applied to the distant view of the scene (Figure 2). As indicated in Figure 6, the saliency patterns in the $\hat{\psi}_\omega$-image were too small to yield the fixed points passing through the halting condition (16). Nevertheless, the comparison of the lower subwindows in Figures 6 and 7 demonstrates that the *as-is* primary in the close-up image can be well estimated by using distant views of the scene images. This implies that even for too distant landmarks, the subtle chromatic complexity can be evaluated to preset the fractal articulation process *prior to* physical approach.

6. **Perceptual Equivalence.** To localize the landmark, the fractal model ν on the saliency pattern $\mathfrak{O}_{\hat{\pi}_i}$ should be consistent with the perspective of the scene images. Let a segment of the future trajectory be downloaded as shown in Figure 10 and consider the consistency of the fractal code ν with the perspective induced by the segment image. As the basis of the scale analysis, we can apply the subcorrelation (1) to the entire image plane and match the generic perspective model to design a consistent fractal model spanning the roadway area. In Figure 3, the existence of a connected roadway area is verified via the generation of a closed graph.



(a) Closed graph by ν_d (b) Closed graph by ν

FIGURE 11. Finite invariance test on capturing probability

Such a computational verification can be extended to the detection of a connected plane supporting the fractal code ν. Figure 11 displays an example of such computational verification. In this case, the scale shift rule along the segment vector is matched with the scale information $\hat{\sigma}_\omega$ to identify a fractal model ν_d spanning an open space on the roadway area as indicated (a) where the probability distribution $\varphi_\rho(\omega|\nu_d)$ for capturing associated fractal attractors is displayed as a smooth field with local maxima $\tilde{\Theta}$ indicated by small blue dots. Due to the sub-correlation of $\hat{\sigma}_\omega$ with local scale distribution, the distribution $\varphi_\rho(\omega|\nu_d)$ maintains the information on boundary objects. Through the detection of the discrete information $\tilde{\Theta}$, we can observe the noise patterns spanning a connected vertical surface of actual objects as well. As indicated in Figure 11(b), we can verify the 2D connectedness of the saliency pattern through the finite invariance test of the fractal model ν on the discrete information $\tilde{\Theta}$. Adding to it, the saliency pattern can be associated with the ground model ν_d on the same information $\tilde{\Theta}$ as shown in Figure 11(a). By combining these results, the saliency pattern is localized as a surface of a boundary object *prior to* range data acquisition.

In the case of the close-up image oriented by the future segment as shown in Figure 10, we can verify a perceptual consistency of the multi-fractal articulation as illustrated in Figure 11(b). The object mode ν designed on the saliency pattern $\mathfrak{O}_{\hat{\pi}_i}$ generates a closed graph on $\tilde{\Theta}$ under the condition of the ground model ν_d. This implies that the landmark model ν is essentially consistent with the roadway model ν_d inducing the perception of the perspective underlying the scene image. As a result of such a perceptual equivalence, we can visualize an object at a saliency pattern within the context of the ground-object structure as illustrated in Figure 12.

Figures 13 and 14 show the results of other experiments. In these experiments, the $\hat{\psi}_\omega$-filter was applied to the scene image (a) to extract saliency patterns as indicated in (b). Associated fractal models were verified on the scale space information and visualized in the ground-object structure as illustrated in (c) and (d), respectively. The experimental results demonstrate that the environmental saliency provides an effective cue for analyzing even ill-conditioned scenes where landmark objects are observed as non-dominant patterns. The post is a rather 'low-keyed' object in a night view as shown in Figure 13. A warning color of vehicles sometimes yields a smaller image than attractive distractions as displayed in Figure 14.

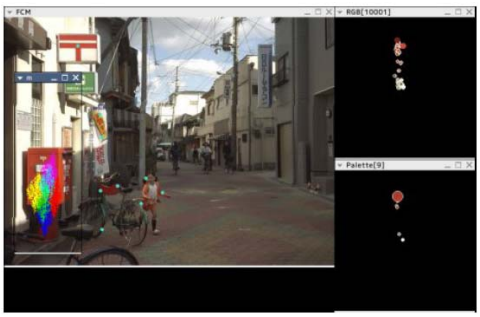
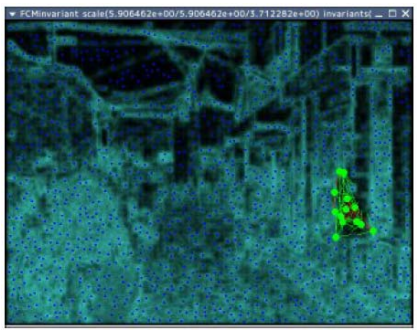


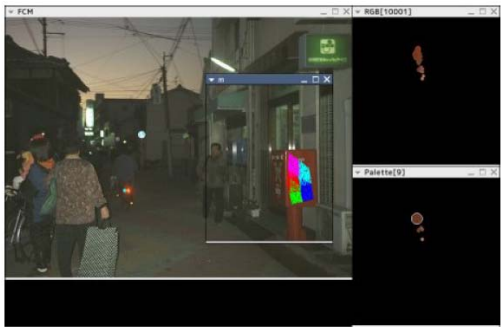
FIGURE 12. Contextual visualization

(a) Scene image

(b) Saliency distribution via $\hat{\psi}_\omega$-filtering



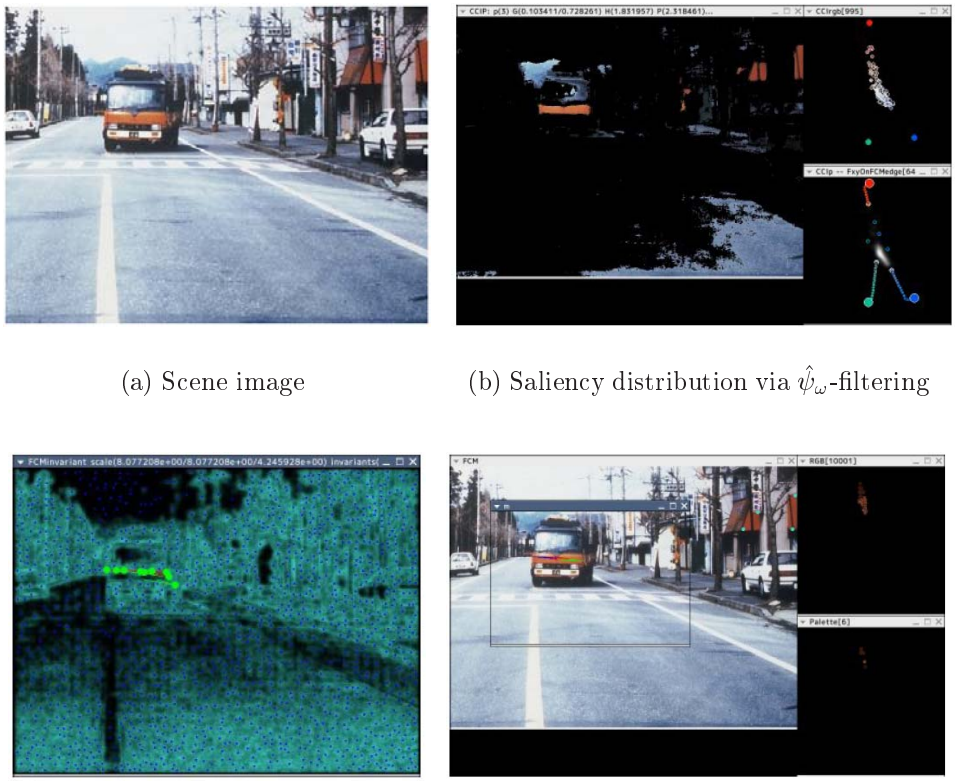
(c) Perceptual equivalence



(d) Contextual visualization

FIGURE 13. Scale-chromatic saliency analysis (street view at night)

By using $\hat{\psi}_\omega$-filtering, we can control the focus of on-vehicle vision within naturally complex scenes. The $\hat{\psi}_\omega f_\omega^{\mathtt{RGB}}$ distribution can be well articulated into a set of fractal attractor. The fractal model can be matched with the perspective underlying the scene image. Through such a 2.5D screening, the fractal model yields an effective cue to the detection of landmark objects by machine- and inherent perception processes. The significance of the $\hat{\psi}_\omega$-filtering is summarized in Table 1, where the reduction of the computational complexity is evaluated in terms of relative entropy $d\mathcal{S}_{(.)} = \mathcal{S}_\emptyset - \mathcal{S}_{(.)}$; $\mathcal{S}_\emptyset$ and $\mathcal{S}_G$ designate the Shannon's entropy with respect to the uniform distribution and the gray level distribution f_ω, respectively; $\mathcal{S}_\mathcal{H}$ and $\mathcal{S}_{\hat{\mathcal{H}}}$ are computed by using normalized version



(a) Scene image (b) Saliency distribution via $\hat{\psi}_\omega$-filtering

(c) Perceptual equivalence (d) Contextual visualization

FIGURE 14. Scale-chromatic saliency analysis (industrial park)

of saliency index as follows:

$$\mathcal{S}_{\mathcal{H}} = -\int_\Omega \left(\frac{\psi_\omega}{C_\psi}\right) \log\left(\frac{\psi_\omega}{C_\psi}\right) d\omega, \qquad C_\psi = \int_\Omega \psi_\omega d\omega,$$
$$\mathcal{S}_{\hat{\mathcal{H}}} = -\int_\Omega \left(\frac{\hat{\psi}_\omega}{C_{\hat{\psi}}}\right) \log\left(\frac{\hat{\psi}_\omega}{C_{\hat{\psi}}}\right) d\omega, \qquad C_{\hat{\psi}} = \int_\Omega \hat{\psi}_\omega d\omega.$$

By definition, the unpredictability of the patterns to be detected is indexed in terms of $e^{-d\mathcal{S}_{(\cdot)}}$; Table 1 implies that the $\hat{\psi}_\omega$-filter can reduce the length of decision makings in many natural scenes to $e^{-d\mathcal{S}_{\hat{\mathcal{H}}}} < 1/7$ of a random search. This amount to 1/6 of pattern detections within the brightness distribution f_ω. In many natural scenes, the *as-is* primary yields a more efficient saliency indexing. Following the experimental results, the complexity of pattern detection remains $10 \sim 30\%$ of conventional RGB-based saliency analysis. As indicated in Table 1, the complexity of landmark selection by inherent and/or machine vision is significantly reduced by the generation and/or indication of the $\hat{\psi}_\omega f_\omega^{\mathtt{RGB}}$ image. Noticing the evaluation $3 \leq \|\hat{\Pi}\| \leq 3+2$, the reduction is considered as the result of the locally Gaussian indexing of the chromatic diversity (2). The robustness of the split and adaptation model (6)- (9) implies the validity of the fractal coding approach to the identification of the chromatic perception process.

7. **Concluding Remarks.** A multi-fractal coding was applied for 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 can be selected to discriminate

TABLE 1. Complexity reduction via ψ_ω-filtering

scene	$d\mathcal{S}_G$	$d\mathcal{S}_{\mathcal{H}}$	$d\mathcal{S}_{\hat{\mathcal{H}}}$	$\|\|\hat{\Pi}\|\|$
shopping street	0.131512	0.995172	2.032901	4
post office	0.199979	2.016679	2.884059	5
street view at night	0.148436	1.481229	3.189055	3
industrial park	0.103411	1.831957	2.059536	3
3d puzzle	0.196245	0.616455	1.266864	4
complexity reduction	$\sim 1/1.2$	$1/1.8 \sim 1/7.4$	$1/3.3 \sim 1/24$	

the image of landmark objects in a noisy background. Supported by the consistency with the ground-object structure induced by the scale complexity, the multi-fractal model of the object images can be exploited to indicate various types of maneuvering contexts arising in the scenes. By using $\hat{\psi}_\omega$-filtering, the computational complexity of landmark detection can be significantly reduced.

REFERENCES

[1] M. F. Barnsley, *Superfractals*, Cambridge University Press, Cambridge, UK, 2006.

[2] M. Ebner, *Color Constancy*, John Wiley and Sons, West Sussex, UK, 2007.

[3] S. Edelkamp, S. Jabbar and T. Willhalm, Geometric travel planning, *IEEE Trans. on Intelligent Transportation Systems*, vol.6, no.1, pp.5-16, 2005.

[4] I. Fujita, K. Tanaka, M. Ito and K. Cheng, Columns for visual features of objects in monkey inferotemporal cortex, *Nature*, vol.360, no.26, pp.343-346, 1992.

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

[6] C. M. Hagelhall, T. Purcell and R. Taylor, Fractal dimension of landscape silhouette as a predictor for landscape preference, *Journal of Emvironmental Psychology*, vol.24, pp.247-255, 2004.

[7] L. Itti, C. Kock and E. Niebur, A model of saliency-based visual attention for rapid scene analysis, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.20, no.11, pp.1254-1259, 1998.

[8] F. A. Jafar, Y. Suzuki, Y. Tateno, K. Yokota and T. Matsuoka, An environmental visual features based navigation method for autonomous mobile robots, *International Journal of Innovative Computing, Information and Control*, vol.7, no.3, pp.1341-1355, 2011.

[9] A. G. Jones and C. J. Taylor, Robust shape from shading, *Image and Vision Computing*, 1994.

[10] K. Kamejima, Generic representation of self-similarity via structure sensitive sampling of noisy imagery, *Electronic Notes in Theoretical Computer Science*, vol.46, pp.225-244, 2001.

[11] K. Kamejima, Laplacian-gaussian sub-correlation analysis for scale space imaging, *International Journal of Innovative Computing, Information and Control*, vol.1, no.3, pp.381-399, 2005.

[12] K. Kamejima, Randomness-based scale-chromatic image analysis for interactive mapping on satellite-roadway-vehicle network, *Journal of Systemics, Cybernetics and Informatics*, vol.5, no.4, pp.78-86, 2007.

[13] K. Kamejima, Chromatic information adaptation for complexity-based integration of multi-viewpoint imagery – A new approach to cooperative perception in naturally complex scene –, *International Journal of Innovative Computing, Information and Control*, vol.4, no.1, pp.109-126, 2008.

[14] K. Kamejima, Object-ground separation via stochastic depth analysis – Anticipative roadway pattern modeling in naturally complex scene –, *International Journal of Innovative Computing, Information and Control*, vol.5, no.1, pp.183-199, 2009.

[15] K. Kamejima, Anticipative coding and *in-situ* adaptation of maneuvering affordance in a naturally complex scene, in *Advances in Human-Robot Interaction*, V. A. Kulyukin (ed.), Vukovar, Croatia, In-Teh, 2010.

[16] K. Kamejima, Environmental saliency: Stochastic representation and dynamic perception, *Proc. of the 14th World Multi-Conference on Systemics, Cybernetics and Informatics*, Orlando, FL, USA, vol.2, pp.92-97, 2010.

[17] V. Kroumov, J. Yu and K. Shibayama, 3D path planning for mobile robots using simulated annealing neural neywork, *International Journal of Innovative Computing, Information and Control*, vol.6, no.7, pp.2885-2899, 2010.
[18] S. Kwak, B. Ko and H. Byun, Automatic salient-object extraction using the contrast map and salient points, *Proc. of Pacific-Rim Conference on Multimedia*, vol.2, pp.138-145, 2004.
[19] C.-H. Lee and M.-F. Lin, Content-based image retrieval for three-dimensional trademarks, *International Journal of Innovative Computing, Information and Control*, vol.6, no.9, pp.3929-3941, 2010.
[20] A. Levin, A. Rav-Acha and D. Lischinski, Spectral matting, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.30, no.10, pp.1699-1712, 2008.
[21] D. G. Lowe, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, vol.60, no.2, pp.91-110, 2004.
[22] S. Lyu and E. P. Simoncelli, Modeling multiscale subbands of photographic images with fields of gaussian scale mixtures, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.31, no.4, pp.693-706, 2009.
[23] D. Marr, *Vision – A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman and Company, San Francisco, CA, 1982.
[24] P. Milanese, S. Gil and T. Pun, Attentive mechanisms for dynamic and static scene analysis, *Optical Eng.*, vol.34, no.8, pp.2428-2434, 1995.
[25] Y. Miyashita and H. S. Chang, Neural correlate of pictorial short-term memory in the primate temporal cortex, *Nature*, vol.331, no.7, pp.568-570, 1988.
[26] 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*, vol.13, no.11, pp.4700-4719, 1993.
[27] Ü. Özgner and C. Stiller, Systems for safety and autonomous behavior in cars: The DARPA grand challenge experience, *Proc. of the IEEE*, vol.95, no.2, pp.397-411, 2007.
[28] H. Palangi and M. H. Refan, Error reduction of a low cost GPS receiver for kinematic applications based on a new Kalman filtering algorithm, *International Journal of Innovative Computing, Information and Control*, vol.6, no.8, pp.3775-3786, 2010.
[29] A. Parker, *In the Blink of an Eye*, The Free Press, London, UK, 2003.
[30] H.-S. Tan and J. Huang, DGPS-based vehicle-to-vehicle cooperative collision warning: Engineering feasibility viewpoints, *IEEE Trans. on Intelligent Transportation Systems*, vol.7, no.4, pp.415-428, 2006.
[31] C. Urmson, C. Baker, J. Dolan, P. Rybski, B. Salesky, W. Whittaker, D. Ferguson and M. Darms, Autonomous driving in traffic: Boss and the urban challenge, *AI Magazine*, vol.30, no.2, pp.17-28, 2009.
[32] J. Wang, S. Schroedl, K. Mezger, R. Ortloff, A. Joos and T. Passegger, Lane keeping based on location technology, *IEEE Trans. on Intelligent Transportation Systems*, vol.6, no.3, pp.351-356, 2005.
[33] C.-P. Young, B.-R. Chang, H.-F. Tsai, R.-Y. Fang and J.-J. Lin, Vehicle collision avoidance system using embedded hybrid intelligent prediction based on vision/GPS sensing, *International Journal of Innovative Computing, Information and Control*, vol.5, no.12(A), pp.4453-4468, 2009.