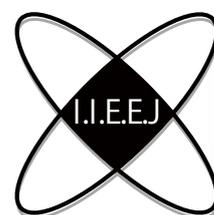


IIEEJ Transactions on Image Electronics and Visual Computing

Special Issue on CG & Image Processing Technologies, for Automation,
Labor Saving and Empowerment Part II

Vol. 9, No. 1 2021



The Institute of Image Electronics Engineers of Japan

Editor in Chief

Mei KODAMA (Hiroshima University)

Vice Editors in Chief

Osamu UCHIDA (Tokai University)

Naoki KOBAYASHI (Saitama Medical University)

Yuriko TAKESHIMA (Tokyo University of Technology)

Advisory Board

Yasuhiko YASUDA (Waseda University Emeritus)

Hideyoshi TOMINAGA (Waseda University Emeritus)

Kazumi KOMIYA (Kanagawa Institute of Technology)

Fumitaka ONO (Tokyo Polytechnic University Emeritus)

Yoshinori HATORI (Tokyo Institute of Technology)

Mitsuji MATSUMOTO (Waseda University Emeritus)

Kiyoshi TANAKA (Shinshu University)

Shigeo KATO (Utsunomiya University Emeritus)

Editors

Yoshinori ARAI (Tokyo Polytechnic University)

Chee Seng CHAN (University of Malaya)

Naiwala P. CHANDRASIRI (Kogakuin University)

Chinthaka PREMACHANDRA (Shibaura Institute of Technology)

Makoto FUJISAWA (University of Tsukuba)

Issei FUJISHIRO (Keio University)

Kazuhiko HAMAMOTO (Tokai University)

Madoka HASEGAWA (Utsunomiya University)

Ryosuke HIGASHIKATA (Fuji Xerox Co., Ltd.)

Tomokazu ISHIKAWA (Toyo University)

Masahiro ISHIKAWA (Saitama Medical University)

Naoto KAWAMURA (Canon OB)

Shunichi KIMURA (Fuji Xerox Co., Ltd.)

Shoji KURAKAKE (NTT DOCOMO)

Kazuto KAMIKURA (Tokyo Polytechnic University)

Takashi KANAI (The University of Tokyo)

Tetsuro KUGE (NHK Engineering System, Inc.)

Koji MAKITA (Canon Inc.)

Junichi MATSUNOSHITA (Fuji Xerox Co., Ltd.)

Tomoaki MORIYA (Tokyo Denki University)

Paramesran RAVEENDRAN (University of Malaya)

Kaisei SAKURAI (DWANGO Co., Ltd.)

Koki SATO (Shonan Institute of Technology)

Syuhei SATO (University of Toyama)

Masanori SEKINO (Fuji Xerox Co., Ltd.)

Kazuma SHINODA (Utsunomiya University)

Mikio SHINYA (Toho University)

Shinichi SHIRAKAWA (Aoyama Gakuin University)

Kenichi TANAKA (Nagasaki Institute of Applied Science)

Yukihiro TSUBOSHITA (Fuji Xerox Co., Ltd.)

Daisuke TSUDA (Shinshu University)

Masahiro TOYOURA (University of Yamanashi)

Kazutake UEHIRA (Kanagawa Institute of Technology)

Yuichiro YAMADA (Genesis Commerce Co., Ltd.)

Norimasa YOSHIDA (Nihon University)

Toshihiko WAKAHARA (Fukuoka Institute of Technology OB)

Kok Sheik WONG (Monash University Malaysia)

Reviewer

Hernan AGUIRRE (Shinshu University)

Kenichi ARAKAWA (NTT Advanced Technology Corporation)

Shoichi ARAKI (Panasonic Corporation)

Tomohiko ARIKAWA (NTT Electronics Corporation)

Yue BAO (Tokyo City University)

Nordin BIN RAMLI (MIMOS Berhad)

Yoong Choon CHANG (Multimedia University)

Robin Bing-Yu CHEN (National Taiwan University)

Kiyonari FUKUE (Tokai University)

Mochamad HARIADI (Sepuluh Nopember Institute of Technology)

Masaki HAYASHI (UPPSALA University)

Takahiro HONGU (NEC Engineering Ltd.)

Yuukou HORITA (University of Toyama)

Takayuki ITO (Ochanomizu University)

Masahiro IWAHASHI (Nagaoka University of Technology)

Munetoshi IWAKIRI (National Defense Academy of Japan)

Yuki IGARASHI (Meiji University)

Yoshihiro KANAMORI (University of Tsukuba)

Shun-ichi KANEKO (Hokkaido University)

Yousun KANG (Tokyo Polytechnic University)

Pizzanu KANONGCHAIYOS (Chulalongkorn University)

Hidetoshi KATSUMA (Tama Art University OB)

Masaki KITAGO (Canon Inc.)

Akiyuki KODATE (Tsuda College)

Hideki KOMAGATA (Saitama Medical University)

Yushi KOMACHI (Kokushikan University)

Toshihiro KOMMA (Tokyo Metropolitan University)

Tsuneo KURIHARA (Hitachi, Ltd.)

Toshiharu KUROSAWA (Matsushita Electric Industrial Co., Ltd. OB)

Kazufumi KANEDA (Hiroshima University)

Itaru KANEKO (Tokyo Polytechnic University)

Teck Chaw LING (University of Malaya)

Chu Kiong LOO (University of Malaya) F

Xiaoyang MAO (University of Yamanashi)

Koichi MATSUDA (Iwate Prefectural University)

Makoto MATSUKI (NTT Quaris Corporation OB)

Takeshi MITA (Toshiba Corporation)

Hideki MITSUMINE (NHK Science & Technology Research Laboratories)

Shigeo MORISHIMA (Waseda University)

Kouichi MUTSUURA (Shinsyu University)

Yasuhiro NAKAMURA (National Defense Academy of Japan)

Kazuhiro NOTOMI (Kanagawa Institute of Technology)

Takao ONOYE (Osaka University)

Hidefumi OSAWA (Canon Inc.)

Keat Keong PHANG (University of Malaya)

Fumihiko SAITO (Gifu University)

Takafumi SAITO (Tokyo University of Agriculture and Technology)

Tsuyoshi SAITO (Tokyo Institute of Technology)

Machiko SATO (Tokyo Polytechnic University Emeritus)

Takayoshi SEMASA (Mitsubishi Electric Corp. OB)

Kaoru SEZAKI (The University of Tokyo)

Jun SHIMAMURA (NTT)

Tomoyoshi SHIMOBABA (Chiba University)

Katsuyuki SHINOHARA (Kogakuin University)

Keiichiro SHIRAI (Shinshu University)

Eiji SUGISAKI (N-Design Inc. (Japan), DawnPurple Inc. (Philippines))

Kunihiko TAKANO (Tokyo Metropolitan College of Industrial Technology)

Yoshiki TANAKA (Chukyo Medical Corporation)

Youichi TAKASHIMA (NTT)

Tokiichiro TAKAHASHI (Tokyo Denki University)

Yukinobu TANIGUCHI (NTT)

Nobuji TETSUTANI (Tokyo Denki University)

Hiroyuki TSUJI (Kanagawa Institute of Technology)

Hiroko YABUSHITA (NTT)

Masahiro YANAGIHARA (KDDI R&D Laboratories)

Ryuji YAMAZAKI (Panasonic Corporation)

IIEEJ Office

Osamu UKIGAYA

Rieko FUKUSHIMA

Kyoko HONDA

Contact Information

The Institute of Image Electronics Engineers of Japan (IIEEJ)

3-35-4 101, Arakawa, Arakawa-ku, Tokyo 116-0002, Japan

Tel : +81-3-5615-2893 Fax : +81-3-5615-2894

E-mail : hensyu@iieej.org

<http://www.iieej.org/> (in Japanese)

<http://www.iieej.org/en/> (in English)

<http://www.facebook.com/IIEEJ> (in Japanese)

<http://www.facebook.com/IIEEJ.E> (in English)

**IIEEJ Transactions on
Image Electronics and Visual Computing
Vol.9 No.1 June 2021
CONTENTS**

Special Issue on CG & Image Processing Technologies,for Automation, Labor Saving and Empowerment

- 1** Upon the Special Issue on CG & Image Processing Technologies,for Automation, Labor Saving and Empowerment Masanori SEKINO

Contributed Paper

- 2** Robust Defect Detection for Logotype Printed on Embossed Surfaces by Multiple Paired Pixel Consistency Sheng XIANG, Shun'ichi KANEKO

Short Paper

- 12** Weakly Supervised Logo Detection Using a Dual-Attention Dilated Residual Network Rahul KUMAR JAIN, Yutaro IWAMOTO, Taro WATASUE, Tomohiro NAKAGAWA, Takahiro SATO, Xiang RUAN, Yen-Wei CHEN

Regular Section

Contributed Paper

- 20** Interactive Control of Fire Color based on Physicochemical Combustion Ryosuke UCHIKAWA, Hiroki ISHIZUKA, Sei IKEDA, Yoshihiro KURODA, Shunsuke YOSHIMOTO, Osamu OSHIRO

Announcements

- 33** 3rd Call for Papers (Call for Late Breaking Papers)
- 34** Call for Papers : Special Issue on Image-related Technology for Realizing Immersive Media
- 35** Call for Papers : Special Issue on CG & Image Processing Technologies Supporting and Expanding Human Creativities

Guide for Authors

- 36** Guidance for Paper Submission

Upon the Special Issue on CG & Image Processing Technologies for Automation, Labor Saving and Empowerment

Editor: Masanori SEKINO
(FUJIFILM Business Innovation Corp.)

A declining birthrate and an aging population are a common problem in developed countries. In order to solve the labor shortage, it is necessary not only to expand the workforce but also to improve labor productivity. Therefore, improvement in productivity is expected through automation and labor saving technologies using evolving image recognition technology, increasing information from IoT devices, and cheaper robots.

At the call for submission, we received 12 papers, 8 in Japanese and 4 in English. Concerning English papers, one paper was adopted in December 2020 issue, and two additional papers were adopted and included in this issue. By the way, three Japanese papers were adopted in January 2021 issue and one additional Japanese paper was adopted (not cited as special issue paper) in April 2021 issue. The two papers included in this issue were both related to logo detection though the discussion points are rather different.

In our journal, we are soliciting “system development papers” and “practical papers”, which are the categories of the papers that make it easy to present the outcome of research focusing on the practical and applied aspects of technology. As the technology recruited in this special issue will well match with these paper categories, we hope your further submission of the papers related to this special issue especially in these paper categories.

Finally, I would like to appreciate the reviewers and editors for careful review and all the efforts to improve the quality of papers. Also, I would like to thank to the editorial committee members of IIEEJ and the staff at IIEEJ office for various kinds of support.

Robust Defect Detection for Logotype Printed on Embossed Surfaces by Multiple Paired Pixel Consistency

Sheng XIANG[†](*Student Member*), Shun'ichi KANEKO[†]

[†] Graduate school of information science and technology, Hokkaido University

<Summary> Embossed surfaces are widely used in many electronic products to enhance the appearance and user experience. However, there are inevitably some defects in the production process. They are difficult to detect by traditional detection algorithms because of the irregular textures on the surface, which are easily affected by illumination conditions. To achieve robust defect detection for logotype printed on the embossed surface, we proposed a novel defect detection method called multiple paired pixel consistency (MPPC). Firstly, we propose a consistency measure for high consistency pixel pairs to realize a robust defect-free model. Secondly, based on this model, we design a measure for judging defective pixel or defect-free pixel. Experimental results with some real-world defective images demonstrate that our approach can achieve state-of-the-art accuracy in real industrial applications.

Keywords: defect detection, embossed surface, multiple paired pixel consistency, orientation code

1. Introduction

The appearance of product is very important for customers to make a good expression. However, in the production and assembly process, various defects are sometimes generated on the surface of the product. Therefore, defect detection plays an important role in the quality control (QC) of the manufacturing industry. At present, defect detection is often performed by human inspectors. Manual inspection is reliable for a small number of products. However, today, a large number of products, such as mobile phones and personal computers, are being massively produced and delivered to the market. It is impossible for human beings to inspect each product one by one, and it requires considerable time to perform such huge amount of work. In addition, it is very expensive to hire workers for inspecting a huge number of products. Therefore, human inspection is being replaced by automatic visual inspection systems¹.

There are many defect detection methods have been proposed in the past several years. From the viewpoint of the algorithm, existing visual inspection methods can be classified to four categories². Statistical approaches, which based the spatial information to evaluate textures, are general applied to irregularly texture case. They including histogram properties³, co-occurrence matrices⁴, autocorrelation⁵, and local binary patterns⁶. Struc-

tural approaches, which are usually consider texture to be characterized by texture primitives, and the spatial arrangement of these primitives⁷, are effective for regular textures. However, utilizing structural methods to detect defects has certain drawbacks, i.e., they are not applicable to structures with a low degree of regularity or small defects. Filter-based approaches, which can be divided into spatial-domain, frequency-domain, and joint-domain techniques, apply filter banks to an image for computing the energy of the filter responses, which include the Fourier transform⁸, Gabor transform⁹, and orthogonal wavelet transform¹⁰. They are effective and simple to implement, but are ineffective with images with a complex texture. Model-based approaches, which detect defect by designing or learning a separation model, take into account both spatial texture and statistic information. Such approaches include the Gaussian-Markov random field model¹¹ and the Gaussian mixture model¹² as representatives. Their results have been satisfactory; however, the computational complexity is high, and they are not good at detecting small defects. And recently, deep learning based method have attracted significant attentions and obtained highly satisfactory results¹³. However, it takes a lot of time for training. In addition, it is difficult and impractical to collect and label plenty of samples.

We focus on printed logotypes on a surface embossed

with 3D micro-textures distributed randomly. This type of surface is formed by embossing processes that create mixed miniature convex and concave patterns on the surfaces. And these surfaces have been widely used in many products, because they have an attractive appearance, good feel, and excellent slip resistance. Furthermore, the changes in illumination on such surfaces have significant influence on the appearance of the surface, which makes it difficult to perform defect detection. To overcome this problem, we use the orientation code (OC) [14] as the basis for the algorithm proposed in this study. OC was previously proposed to be used for robust image matching, and its robustness to illumination fluctuations was verified in different scenes. Furthermore, based on this modeling, we propose an algorithm for defect detection mainly in the logotype regions, which any customer maybe sensitive to check the quality of their printing.

The proposed model called multiple paired pixel consistency (MPPC) for defect detection, was briefly introduced in¹⁵⁾. In this paper, we introduce MPPC with some new contents and give more detailed explanations for our work with new experiment results. The contribution of this paper are as follows: (1) this work propose the MPPC model of defect-free to realize the robustness of illumination fluctuations for detecting the defect in logotypes printed on embossed surfaces; (2) a new definition of OC difference is proposed, which is better suited for the detailed statistical design; (3) the parameter setting in practical application is discussed, which can make parameter setting easier for readers to achieve good detection results.

The rest of this paper is organized as follows. Section 2. introduces the working mechanism of the MPPC defect-free model. Section 3. presents defect detection using the MPPC model, and to show the effectiveness of the proposed method some experimental results are given in Section 4.. Finally, some concluding remarks are given in Section 5..

2. MPPC Model for Defect-Free Image

In this section, we first introduce the original version of the orientation code, and then, propose the signed difference between any two codes in preparation for developing a more precise statistical model to represent their differences. In addition, we provide a precise spatial differentiation for calculating the codes with a higher resolution. Finally, we introduce how to construct the MPPC model from defect-free logos.

2.1 Orientation codes

The orientation codes were proposed for applying a filtering to extract the robust features based on only the orientation information involved in gradient vectors from any images. This approach achieves strong robustness to an image representation and can resist the fluctuations in the illumination in the real world because it utilizes only a gradient orientation as a discriminative feature, which represents the orientation angle rather than the strength of the gradient.

Let $I(i, j)$ be the brightness of pixel (i, j) . Its partial derivatives in the horizontal and vertical directions are written as $\nabla I_x = \partial I / \partial x$ and $\nabla I_y = \partial I / \partial y$, respectively. The orientation angle θ can be computed using $\theta = \tan^{-1}(\nabla I_y / \nabla I_x)$ of which the actual orientation is determined after checking the signs of the derivatives, thus making the range of θ be $[0, 2\pi)$. The OC is obtained by quantifying the orientation angle θ into N levels with a constant width $\Delta\theta (= 2\pi/N)$. The OC can be expressed as follows:

$$C_{i,j} = \begin{cases} \left\lfloor \frac{\theta_{i,j}}{\Delta\theta} \right\rfloor & |\nabla I_x| + |\nabla I_y| \geq \Gamma \\ N & \text{otherwise} \end{cases} \quad (1)$$

where Γ is a threshold level for ignoring pixels with low-contrast neighborhoods. The pixels with neighborhoods of sufficient contrast are assigned an OC of $\{0, 1, \dots, N - 1\}$, whereas the ignored pixels are assigned the code N . **Figure 1** shows the code system used in this paper, where $\Gamma = 10$ and $N = 16$.

Here, a new definition of OC difference, which is better suited for the detailed statistical design, is proposed. Compared with the previous definition¹⁴⁾, the definition used here has both positive and negative differences. It can provide a more complete and precise distribution of

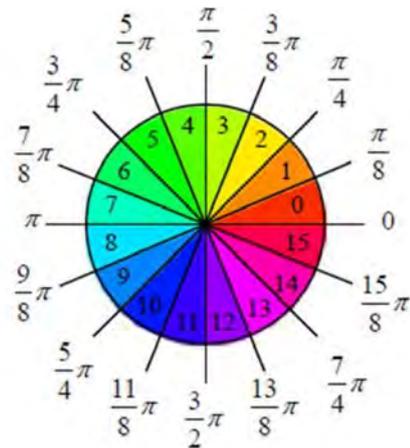


Fig. 1 Sixteen-OC

the differences in OC. In this way, we will obtain a symmetric distribution of the OC difference, which facilitates statistical handling. This is expressed as follows:

$$\Delta(a, b) = \begin{cases} \min\{|a - b|, N - |a - b|\} & (a - b) < -\frac{N}{2}, \\ 0 & 0 \leq (a - b) < \frac{N}{2} \\ -\min\{|a - b|, N - |a - b|\} & \text{otherwise} \end{cases} \quad (2)$$

where a and b represent the OC to be compared or subtracted, for instance, from the target and reference images, respectively, and N shows the invalid-pixel code.

Since we are dealing with objects having very fine textures, and furthermore, we aim to detect extremely small defects during the real factory production in this study, the manufacturing processes continue to advance, and the defects occurring in actual production lines continue to decrease. Yet, the existing Sobel operator cannot obtain enough texture information for fine textures. Therefore, to realize a better spatial sensitivity, rather than greater stability, we proposed a new gradient or differentiation operator for obtaining a higher resolution or definition. In the next equation 3, we define two convolution matrices for the gradient operator in a 2×2 region.

$$G_x = \begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix}, G_y = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix} \quad (3)$$

A pixel of interest in a convolution matrix is defined at the lower-left position in this 2×2 region and may have a weight of -2; the neighboring pixels along the x and y axes then have a weight of +2 whereas the neighboring pairs of pixels have a weight of -1 or +1, allowing a smoothing effect for de-noising.

Figure 2 shows three typical examples used to evaluate the sensitivity by comparing the proposed operator with a Sobel operator using a standard Lena image. On comparing the Figs. 2 (c) and (d), we can see that applying proposed operator results in a thinner edge without increasing noise levels. Then, by observing the partial enlargement image, the proposed operator can extract a more complete texture. For instance, when the Sobel operator is applied in the upper part of the hat, the extracted texture is partially missing. On the other hand, the proposed operator can extract a more complete texture. This shows that the proposed operator is more sensitive to edges.

2.2 MPFC defect-free model

2.2.1 Preparation

As the first step toward modeling the statistical behavior of pixels on a logotype on an embossed surface, we

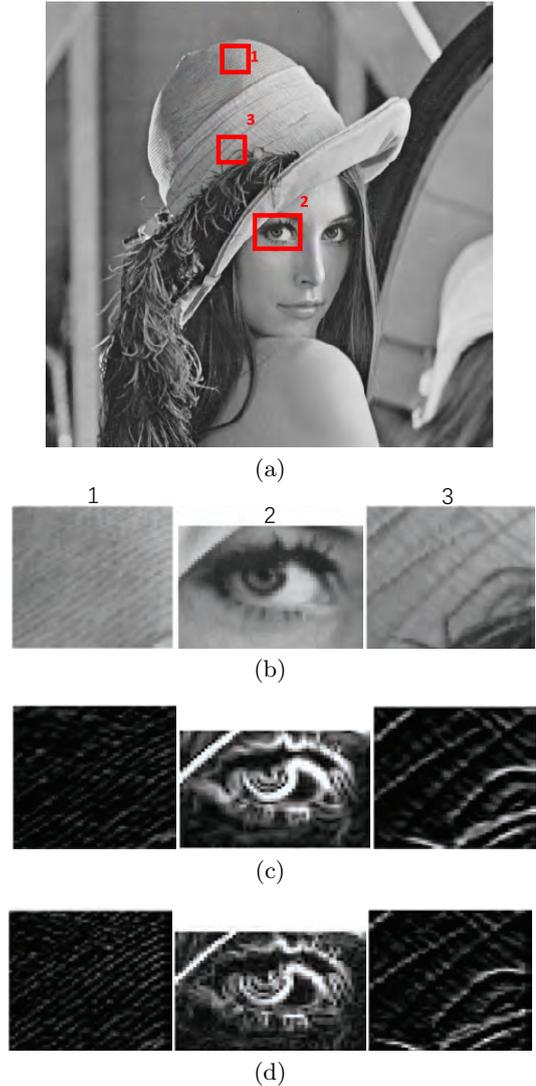


Fig. 2 Examples of applying the proposed and Sobel operators; (a) Lena, (b) magnified parts, (c) Sobel operator, and (d) proposed operator

consider the relationship between any pair of two pixels on the logotype. Here, we aim to investigate what type of relation any pair of pixels in different locations has in the OC difference. In one pair, both pixels are on the logotype, whereas in the other pair one pixel is on the logotype and the other is on the background. **Figure 3** shows two typical examples of OC difference histograms made from a dataset of 160 logotype images. As an example of a typical logotype defect, such as a lack of printing, we can see that one of the logo-background pairs, shown in Fig. 3(b), has a wide variance compared with the logo-logo pairs shown in Fig. 3(a), where the profile has a clear peak as another prominent feature in comparison with the latter; we can also see a symmetric property around the mean or center value.

From this observation and the rareness of the defects occurring in actual factories, we assume that they can ba-

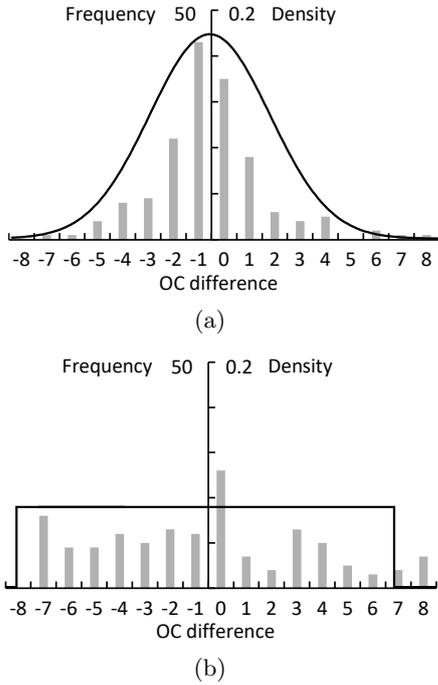


Fig. 3 Two representative types of OC difference in a pair of pixels; (a) logo-logo pair and (b) logo-background pair

sically be fitted by a single Gaussian distribution. Their kurtoses are calculated to find their potential distributions. We found that the kurtosis of the histogram in Fig. 3(a) was 0.01, which is close to 0, making it reasonable to fit or model using a single Gaussian distribution. The histogram can be modeled by the single Gaussian distribution with a mean of -0.56 and a standard deviation of 2.3. By contrast, for a histogram of logo-background pairs, the calculated kurtosis was -1.2, which shows that it can be modeled using a uniform distribution. Fig. 3(b) can be fitted with a uniform distribution, and their mean and variance are -1 and 17.6, respectively. Based on these observations, we ascertain the following significant differences in the behaviors of these pairs: logo-logo pixel pairs may generally have an extremely high correlation owing to the logotype printing, whereas logo-background pairs have no clear correlative relation. We expect the latter type of pair to possibly be used to represent typical defects, such as a lack of printing.

Figure 4 (a) shows a schematic structure of the proposed MPPC model, which can represent one statistical characteristic in the OC difference between two elemental pixels among the pairs. The main idea of this statistical modeling of images is called ‘CP3’, which was previously proposed by Liang¹⁶⁾ for robust background subtraction. We propose an extension of this scheme by introducing the cohesive relationship of OCs in logo-logo pairs de-

finied between each target pixel P on a logotype and the set of supporting pixels, also on the same logotype, which should be selected to achieve a higher consistency or correlation with the target pixel. In other words, similar trends of change as the target pixel are shown, for which we can make a statistical model by fitting a single Gaussian distribution to the OC difference histogram of these pairs with high consistency.

2.2.2 Consistency based supporting pixel selection

We now consider how to select the supporting pixels from all candidate pixels for a target pixel. For formalization, in this paper, we use capital letters in bold face, such as \mathbf{Q} , to represent any set, capital letters in normal font to represent any pixel, and lowercase letters to show any OCs of the pixels.

For an arbitrary logo-logo pixel pair (P, Q) , we have two sets of OC sequences at the same positions in all K training images as follows:

$$\mathbf{P} = \{p_1, p_2, \dots, p_K\} \quad (4)$$

and

$$\mathbf{Q} = \{q_1, q_2, \dots, q_K\}, \quad (5)$$

where K is the total number of training sample images, as shown in Fig. 4(a). And the expected values and variances over \mathbf{P} and \mathbf{Q} are defined as follows:

$$\bar{p} = \frac{1}{K} \sum_{k=1}^K p_k, \bar{q} = \frac{1}{K} \sum_{k=1}^K q_k \quad (6)$$

$$\sigma_{\mathbf{P}}^2 = \frac{1}{K} \sum_{k=1}^K (p_k - \bar{p})^2, \sigma_{\mathbf{Q}}^2 = \frac{1}{K} \sum_{k=1}^K (q_k - \bar{q})^2 \quad (7)$$

The covariance between \mathbf{P} and \mathbf{Q} is defined as follows:

$$C_{\mathbf{P}, \mathbf{Q}} = \frac{1}{K} \sum_{k=1}^K (p_k - \bar{p})(q_k - \bar{q}) \quad (8)$$

If $C_{\mathbf{P}, \mathbf{Q}} > 0$, a consistency or co-occurrence probability occurs, and to measure the consistency quantitatively, we use the following Pearson’s product-moment correlation coefficient:

$$\gamma_{\mathbf{P}, \mathbf{Q}} = \frac{C_{\mathbf{P}, \mathbf{Q}}}{\sigma_{\mathbf{P}} \cdot \sigma_{\mathbf{Q}}} \quad (9)$$

where $\sigma_{\mathbf{P}}$ and $\sigma_{\mathbf{Q}}$ are the standard deviations of \mathbf{P} and \mathbf{Q} , respectively.

For all pixels $P(u, v)$ at the position (u, v) over the full regions covering the logotype, we may have $(M - 1)$ candidate pixels in the same logotype, where M properly

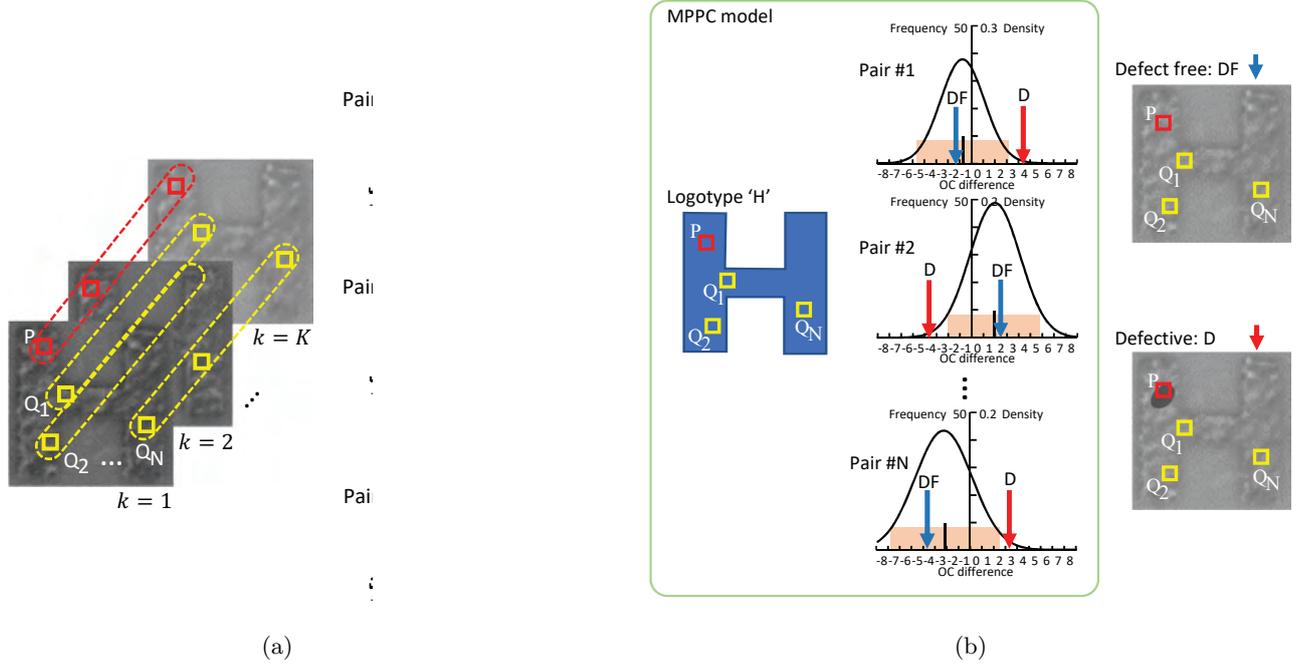


Fig. 4 Scheme of MPPC; (a) MPPC modeling and (b) defect detection based on MPPC model

defines the total size of the logotype in pixels. From these candidates, we can select $N (< M)$ supporting pixels in descending order of the value of $\gamma_{\mathbf{P}, \mathbf{Q}}$. The set of N supporting pixels is as follows:

$$\mathcal{Q} = \left\{ Q_i(u_i, v_i) \mid \gamma_{\mathbf{P}, Q_i} \geq \gamma_{\mathbf{P}, Q_{i+1}} \right\}_{i=1,2,\dots,N} \quad (10)$$

And then record the position u_i, v_i of supporting pixels.

2.2.3 Building defect-free model

In this subsection, we discuss how to build defect-free model based on single Gaussian model. We assume that each supporting pixel Q_i maintains a bivariate OC difference with the target pixel P as follows:

$$\Delta(p, q_i) \sim N(\mu_i, \sigma_i) \quad (11)$$

where $N(\mu_i, \sigma_i)$ is the Gaussian distribution with a mean μ_i and variance σ_i^2 , which are calculated from the corresponding pixel sets \mathbf{P} and \mathbf{Q} .

After modeling for one target pixel P , the above set of N pairs of four parameters for the position u_i, v_i and the two statistical parameters μ_i and σ_i are recorded in a row of the look-up table (LUT). Through repetitive modeling, the LUT is filled in to include the total set of MPPC models for all pixels in an elemental logotype.

3. Defect Detection by MPPC

We now discuss how to utilize the proposed MPPC model of the relationship between pixel pairs in a defect-free logotype for detecting many types of logotype defects.

Because the MPPC model can represent the statistical behavior of the relation of an individual target pixel to the supporting pixels around it, we utilize it to statistically test whether a target pixel is recognized as a reasonable sample from the distribution registered in the LUT. The scheme for the defect detection algorithm applied is shown in Fig. 4(b). This include two main procedures: (1) evaluate the normal or abnormal state of each pixel pair (P, Q_i) ; (2) estimate target pixel P is a defect-free or defective pixel.

3.1 Evaluation of the state of pixel pair

From these considerations and the investigation described in Section 2., the two features of our MPPC model, spatial sparseness and high consistency in a correlative relation, can be utilized to handle such defects. The former may prevent any defects from occupying P and some supporting pixels simultaneously, and through the use of the latter feature we then expect to have an evaluation scheme to recognize whether P is occupied with any defects.

The next task must be to design a measure for judging defective pixels or defect-free pixels using the MPPC model. We first test each OC difference between the target pixel P and a supporting pixel Q in the MPPC model or the LUT by applying Q for P in the target image. We define the following measure,

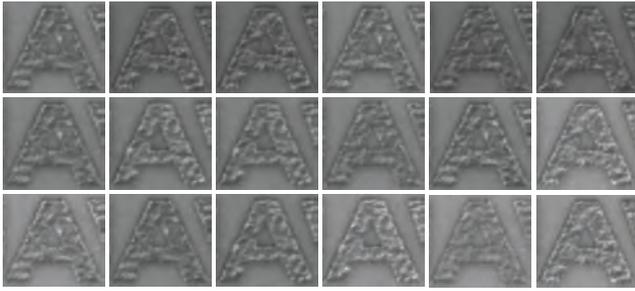


Fig. 5 Some examples of training data

$$\beta_i = \begin{cases} 1 & \|\Delta(p, q_i) - \mu_i\| \geq C \cdot \sigma_i \\ 0 & \text{otherwise} \end{cases}, \quad (12)$$

to identify the normal ($\beta_i = 0$) or abnormal ($\beta_i = 1$) states at the corresponding position defined by the elemental MPPC model, where the constant C is a parameter that can be set from 1 to 3 to define an area for an acceptance probability of 68% through 99.7%.

3.2 Judgement of pixel state

Here, we introduce how to judge the target pixel P is defective or defect free pixel. We use the total sum

$$\xi = \frac{1}{N} \sum_{i=1}^N \beta_i \quad (13)$$

to construct a decision rule for the occupation of P by any defect: $\xi \geq T$, where $T = (\lfloor N/2 \rfloor + 1) / N$ is a threshold is determined from the general majority rule, $\lfloor \cdot \rfloor$ shows the floor function, and N is the total number of supporting pixels.

4. Experiments

4.1 Dataset and evaluation indicators

Our method is tested on a dataset collected from factories. This dataset consists of four sets of real images including four characters ‘A’, ‘E’, ‘I’, and ‘H’, each of which includes 100 defect-free images under three different illumination conditions: dark, normal, and bright.

Figure 5 shows some examples of the training data.

We utilize these defect-free images to train the MPPC model. All the images have a size of approximately 100×100 pixels and a resolution of 15 pixels/mm. The physical size of each character is approximately 7mm. Usually, a defect is less than 10 pixels in length (or less than 1mm) and is usually a hard target for human inspectors to perform visual inspection with their naked eyes. Figure 6 shows some representative examples: (a) micro defects are very small in size and can easily be confused with the texture of logotypes themselves; (b) spots

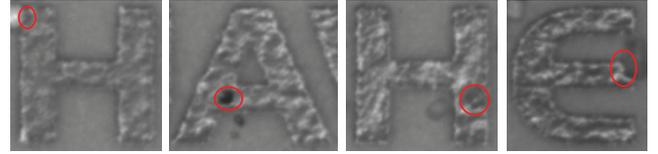


Fig. 6 Representative defect types; From left to right: micro defect, spot, low contrast defect, and missing print

are generally continuous, blocky, and close to circular in shape; (c) low contrast defects are close to logotypes in intensity and have no obvious boundaries; (d) missing print usually occurs at the boundary of the logotype, and the boundary is fuzzy. Through we set the two thresholds of the detection stage as $C = 2.0$ and $T = 0.5$, respectively.

Basically, because the proposed method does not consider the translation or rotation changes of the target image, any original data image should be aligned at the same position and should not have any rotation. In a real factory, it is not very difficult to achieve this requirement due to quality control. In this study, we achieved it through three inspection procedures. The first is that the operator checks all the data. Use a dedicated fixing tool to correct it so that it has a constant position and orientation relative to the camera. Next, we checked their rotation constancy. Finally, the position correlation of all elemental logotype images corresponding to each character is corrected by aligning them with the normalized cross correlation.

To evaluate the methods in pixel level, we utilize three common analysis measurements¹⁷): *Precision* can be seen as a measure of exactness, *Recall* can be seen as a measure of completeness. and *F – measure* as a harmonic average of the *Precision* and *Recall*.

For image-based performance evaluation, the detection rate used was as follows:

$$Detection\ Rate = \frac{N_T}{N_{TD}} \quad (14)$$

where N_T and N_{TD} are the numbers of defective logotypes or images correctly detected and the total number of defective samples, respectively.

4.2 Discussion on parameter setting

In this section, we discuss the influence of some parameters on the defect detection results. They include the parameter C in Eq. (12) and N which is the number of supporting pixels.

The parameter C in Eq. (12) has a direct and important influence on the detection results. As the value of C

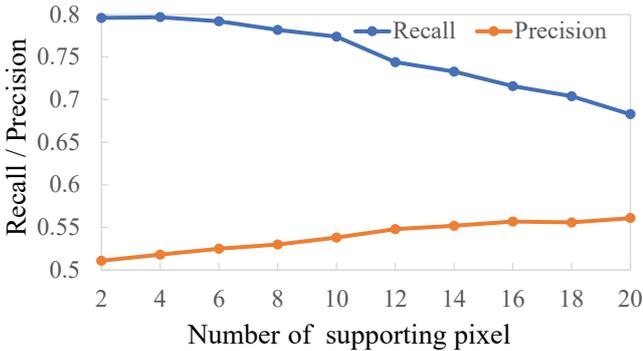
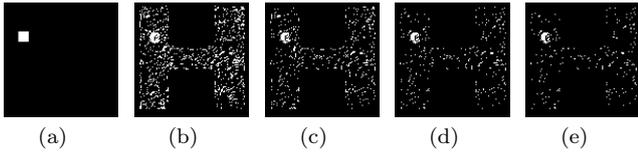


Fig. 8 Quantitative analysis of defect detection

decreases, the integrity of the detection increases; however, the noise will also increase, which reduces precision. In defect detection, the integrity of the detection results is very important, even at the expense of a certain precision. Therefore, through experiments, we set C as 2, and to improve the precision, we utilized a closing morphological operator of size of 5×5 .

The number of supporting pixels N is important for establishing a reliable model. A larger number of supporting pixels can provide more information, which makes subsequent defect extraction more efficient. However, using too many supporting pixels cause two problems. One is that it can result in over-fitting of the extracted information, which results in more noise in the experimental results, and the other is that it will require more computational time.

A basic experiment is designed to examine the defect detection performance using different numbers N , and to find the optimal value of them. Here we set the values in the range of $[2, 4, \dots, 20]$.

Figure 7 shows some the results. From Fig. 7, it can be seen that as the number increases, the defect extraction integrity decreases to some extent, along with a decline of noise level, and furthermore, when the number is very small, the extraction of the defect performance is rather high, while there is considerable noise such as a lot of false positive pixels have appeared. This may be

caused insufficient training models. When the number is increased, the noise level in the detection result is obviously decreased; however, the extracted part of the defect is not complete. This may be due to the over-fitting of the training model. Figure 8 shows the quantitative analysis of character 'H'. From Fig. 8, we can see that the parameter N mainly affects Recall. As the number of supporting pixels increases, there is a slight decrease in Recall, and a slight increase in Precision. When $N = 2$, the Precision is approximately 0.5. According to the definition of Precision, this result means that the number of false positive pixels (FP) and true positive pixels (TP) are similar, which indicates that a very small number of supporting pixel cannot enable robust detection. When the value of N is increased up to 10, Recall can maintain a relatively high level, while Precision gradually increases. When N is increased beyond the value of 10, as N increases, there is a significant drop in Recall, while Precision is almost unchanged. Therefore, we think that 10 is the optimal value of N in this study.

4.3 Robust to fluctuations in the illumination

Precise control of the actual factory illumination conditions is not easy. Furthermore, it is common in factories to have plural sets or stages of visual inspection, which may undergo changes in the conditions over time. Therefore, the proposed algorithm should be robust to fluctuations in the illumination, for which we have introduced OCs to build the MPPC model. Here, we utilize the images of the same objects under five different illumination conditions for the experiments, as shown in Fig. 9, which were collected from real factories under the consideration of real illumination conditions. There are 60 real defect images.

Figure 9 shows some of the detection results. Despite a severe fluctuation in illumination, we found that the proposed method can detect real defects similarly in size and shape from this figure, which shows the strong robustness of the OCs in the MPPC models. Table 1 shows the pixel-level performance evaluation. From Table 1, we can see that under different illumination conditions, the performance of the defect detection is extremely high and achieves similar results. And even under three different lighting conditions, the detection rate reaches 100%. This experimental result shows that our proposed algorithm is very robust to illumination fluctuations. In addition, by comparing the results based on the Sobel operator and the proposed operator, it can be seen that the proposed

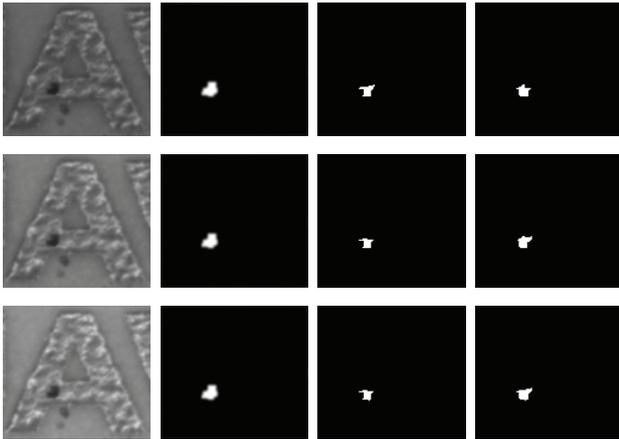


Fig. 9 Some examples of defect detection under different illumination conditions; From left to right: the test images, their ground truth images, the results based on Sobel operator, and the results based on the proposed operator; From top to bottom: images from dark to bright illumination conditions

Table 1 Pixel-level performance of defect detection under different illumination conditions

| Illumination condition | Operator | Recall | Precision | F-measure |
|------------------------|----------|--------|-----------|-----------|
| Dark | Sobel | 0.73 | 0.97 | 0.83 |
| | Proposed | 0.86 | 0.97 | 0.91 |
| Normal | Sobel | 0.72 | 0.97 | 0.83 |
| | Proposed | 0.88 | 0.98 | 0.93 |
| Bright | Sobel | 0.0.72 | 0.97 | 0.83 |
| | Proposed | 0.88 | 0.97 | 0.92 |

operator is capable of detecting more complete defects and is more sensitive to defects.

4.4 Different characters

In this section, we consider the performance of defect detection for different printed characters. We used four cases of real defects in different positions for ‘A’, four for ‘E’, three for ‘I’, and four for ‘H’, respectively. For the logotype ‘T’, we need an additive explanation because the size is relatively small. On the one hand, we cannot find an adequate number of stable supporting pixels, which will impact the effectiveness of our algorithm. On the other hand, the small size may cause the selected supporting pixels to become too clustered, and thus the supporting pixels are easily affected by each other. Therefore, we extend the area to include a part of the neighboring ‘E’, as shown in the third column of **Fig. 10**. This figure shows some examples of defect detection results for each character, We can see that the proposed method can locate most parts of the defect area. After filtering, the over-detection noise is mostly removed and some holes are filled in. In this paper, we focus on the logotype area and

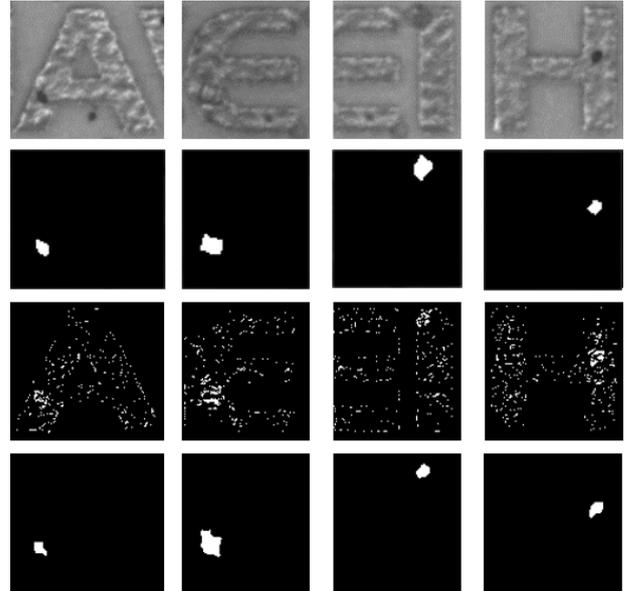


Fig. 10 Some examples of defect detection results for different characters; From the first row: test images, their ground truth images, the detected results, and the filtered results

Table 2 Pixel-level based performance of defect detection for different characters

| Character | Filter | Recall | Precision | F-measure |
|-----------|--------|--------|-----------|-----------|
| A | OFF | 0.65 | 0.62 | 0.64 |
| | ON | 0.88 | 0.96 | 0.92 |
| E | OFF | 0.66 | 0.62 | 0.64 |
| | ON | 0.87 | 0.89 | 0.88 |
| I | OFF | 0.58 | 0.6 | 0.59 |
| | ON | 0.73 | 0.95 | 0.83 |
| H | OFF | 0.62 | 0.65 | 0.63 |
| | ON | 0.83 | 0.97 | 0.90 |

excluded the background area (unprinted area). However, in Fig. 10, the first and third columns contain some defects in the background area. To solve these problems, we will extend our algorithm to include defect detection in the background area in the future.

Table 2 shows a pixel-level based performance evaluation. After filtering, the recall and precision improved substantially, and the F-measure reached approximately 0.85. This shows that our method has achieved good results. For the image-based evaluation, all images with defects were successfully identified, demonstrating the stability of our method.

4.5 Comparison experiment

Many researchers carrying out defect detection of surfaces have commonly examined steel^{18),19)}, textile^{20),21)}, and wood²²⁾. For the visual inspection of printed patterns, examinations have generally been performed using paper materials and pharmaceutical capsules¹⁾. However, to the best of our knowledge, few studies have inves-

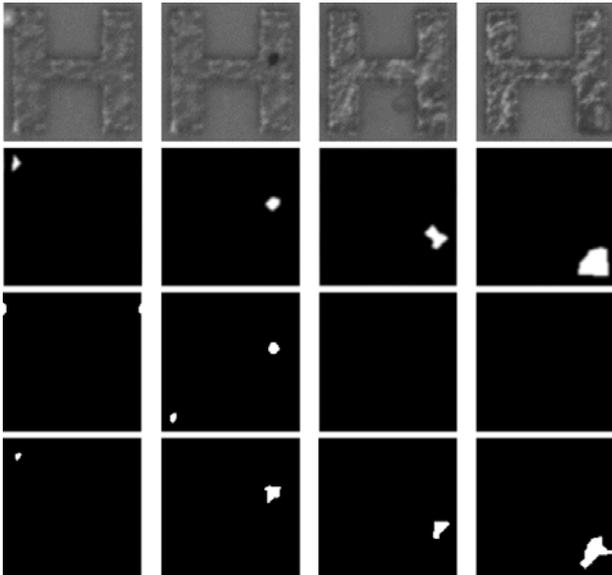


Fig. 11 Some examples of detection results; From top to bottom: test images, their ground truth images, and the detected results by PHOT and the proposed method

Table 3 Pixel-level based performance of defect detection for comparison

| | Recall | Precision | F-measure |
|----------|--------|-----------|-----------|
| PHOT | 0.53 | 0.87 | 0.66 |
| Proposed | 0.83 | 0.97 | 0.90 |

tigated defect detection for logotypes on embossed surfaces. To verify the effectiveness of the proposed method, we compared it with a phase only transform (PHOT) for surface defect detection²³). It has high effectiveness in detecting defects on irregularly textured surfaces, and the objects they deal with are somewhat comparable to the materials mentioned in this paper. Furthermore, the authors disclosed their source codes and the settings of the parameters. Herein, we utilize 40 real defect images for comparison experiments. The implementation was conducted using the source code disclosed, and the parameters selected were those recommended by the respective authors.

Figure 11 shows some representative results. We can see that around half of the defects were detected by PHOT, which is roughly consistent with the ground truth, whereas the proposed method could detect all of the defects. For high-contrast defects, PHOT could obtain a rather good detection performance. However, for low-contrast defects, it did not achieve a good performance. **Table 3** shows the pixel-level based performance evaluation. From this, we can see that our method can achieve good performance. And for image-based evaluation, the PHOT can successfully detect approximately

70% of the defects, whereas the proposed method can successfully detect all defects.

5. Conclusions

In this paper, we proposed a novel defect-free model named as MPPC, for defect detection against illumination fluctuations. Firstly, we defined a consistency measure for high consistency pixel pair to realize a robust defect-free model. In this model, OC difference definition could be combined for more precise modeling. Then, based on this model, a new defect detection scheme was proposed for finding logotype defects. The experimental results for real defect images demonstrate the effectiveness of proposed method. In this study, we focused on the printed logotype, however, the extension of algorithm to perform defect detection in both logotype area and background must be planned. Furthermore, we will attempt to design schemes to identify and classify different defect types to improve the effectiveness of QC in manufacturing production lines.

References

- 1) A. Mehle, M. Bukovec, B. Likar, D. Tomažević: "Print Registration for Automated Visual Inspection of Transparent Pharmaceutical Capsules", *Machine Vision and Applications*, Vol.27, No.7, pp.1087–1102 (2016).
- 2) X. Xie: "A Review of Recent Advances in Surface Defect Detection Using Texture Analysis Techniques", *ELCVIA: Electronic Letters on Computer Vision and Image Analysis*, Vol.7, No.3, pp.1–22 (2008).
- 3) Z. Wakaf, H. A. Jalab: "Defect Detection Based on Extreme Edge of Defective Region Histogram", *Journal of King Saud University-Computer and Information Sciences*, Vol.30, No.1, pp.33–40 (2018).
- 4) A. Latif-Amet, A. Ertüzün, A. Erçil: "An Efficient Method for Texture Defect Detection: Sub-Band Domain Co-Occurrence Matrices", *Image and Vision Computing*, Vol.18, No.6, pp.543–553 (2000).
- 5) D. Zhu, R. Pan, W. Gao, J. Zhang: "Yarn-Dyed Fabric Defect Detection Based on Autocorrelation Function and GLCM", *Autex Research Journal*, Vol.15, No.3, pp.226–232 (2015).
- 6) S. Fekri-Ershad, F. Tajeripour: "Multi-Resolution and Noise-Resistant Surface Defect Detection Approach Using New Version of Local Binary Patterns", *Applied Artificial Intelligence*, Vol.31, No.5, pp 395–410 (2017).
- 7) F. M. Vilmrotter, R. Nevatia, K. E. Price: "Structural Analysis of Natural Textures", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.1, pp.76–89 (1986).
- 8) K. Sakhare, A. Kulkarni, M. Kumbhakarn, N. Kare: "Spectral and Spatial Domain Approach for Fabric Defect Detection and Classification", *Proc. of IEEE 2015 International Conference on Industrial Instrumentation and Control (ICIC)*, pp.640–644 (2015)
- 9) J. Jing, S. Chen, P. Li: "Automatic Defect Detection of Patterned Fabric via Combining the Optimal Gabor Filter and Golden Image Subtraction", *Journal of Fiber Bioengineering and Informatics*, Vol.8, No.2, pp.229–239 (2015).

- 10) Z. Wen, J. Cao, X. Liu, S. Ying: "Fabric Defects Detection Using Adaptive Wavelets", International Journal of Clothing Science and Technology (2014).
- 11) Y. Xiaobo: "Fabric Defect Detection of Statistic Aberration Feature Based on GMRF Model", Journal of Textile Research, Vol.4, p.26 (2013).
- 12) M. Li, S. Cui, Z. Xie: "Application of Gaussian Mixture Model on Defect Detection of Print Fabric", Journal of Textile Research, Vol.36, No.8, pp.94–98 (2015).
- 13) Y. Li, D. Zhang, D. Lee: "Automatic Fabric Defect Detection with a Wide-and-Compact Network", Neurocomputing, Vol.329, pp.329–338 (2019).
- 14) F. Ullah, S. Kaneko: "Using Orientation Codes for Rotation-Invariant Template Matching", Pattern Recognition, Vol.37, No.2, pp.201–209 (2004).
- 15) S. Xiang, Y. Yan, H. Asano, S. Kaneko: "Robust Printing Defect Detection on 3D Textured Surfaces by Multiple Paired Pixel Consistency of Orientation Codes", Proc. of IEEE 2018 12th France-Japan and 10th Europe-Asia Congress on Mechatronics, pp.373–378 (2018).
- 16) D. Liang, M. Hashimoto, K. Iwata, X. Zhao: "Co-Occurrence Probability-Based Pixel Pairs Background Model for Robust Object Detection in Dynamic Scenes", Pattern Recognition, Vol.48, No.4, pp.1374–1390 (2015).
- 17) H. Y. Ngan, G. K. Pang, N. H. Yung: "Automated Fabric Defect Detection a Review", Image and Vision Computing, Vol.29, No.7, pp.442–458 (2011).
- 18) M. Selvi, D. Jenefa: "Automated Defect Detection of Steel Surface Using Neural Network Classifier with Co-occurrence Features", International Journal, Vol.4, No.3 (2014).
- 19) S. Zhou, Y. Chen, D. Zhang, J. Xie, Y. Zhou: "Classification of Surface Defects on Steel Sheet Using Convolutional Neural Networks", Materials and Technology, Vol.51, No.1, pp.123–131 (2017).
- 20) P. Li, J. Liang, X. Shen, M. Zhao, L. Sui: "Textile Fabric Defect Detection Based on Low-Rank Representation", Multimedia Tools and Applications, Vol.78, No.1, pp.99–124 (2019).
- 21) H. Abdellah, R. Ahmed, O. Slimane: "Defect Detection and Identification in Textile Fabric by SVM Method", IOSR Journal of Engineering, Vol.4, No.12, pp.69–77 (2014).
- 22) Y. Zhang, C. Xu, C. Li, H. Yu, J. Cao: "Wood Defect Detection Method with PCA Feature Fusion and Compressed Sensing", Journal of Forestry Research, Vol.26, No.3, pp.745–751 (2015).
- 23) D. Aiger, T. Hugues: "The Phase Only Transform for Unsupervised Surface Defect Detection", Proc. of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.295–302 (2010).



Sheng XIANG (*Student Member*)

He received M.S. degree and Ph.D. degree in system science and informatics from Hokkaido University, Sapporo, Japan, in 2018 and 2021, respectively. His research interests include machine vision and its industrial applications.



Shun'ichi KANEKO

He received the B.S. degree in precision engineering and the M.S. degree in information engineering from Hokkaido University, Japan, in 1978 and 1980, respectively, and then the Ph.D degree in systems engineering from the University of Tokyo, Japan, in 1990. He had been a research assistant of the Department of Computer Science since 1980 to 1991, an associate professor of the Department of Electronic Engineering since 1991 to 1995, and an associate professor of the Department of Bio-application and Systems Engineering since 1995 to 1996, in Tokyo University of Agriculture and Technology, Japan. He had been a professor of the Laboratory of System Sensing and Control in the Graduate School of Information Science and Technology in Hokkaido University, and then retired in March 2020. He received many awards including the Best Paper Award 2014 from the British Physics Society: IOP. His research interest includes machine vision, image sensing and understanding, robust image registration. He is a member of IEICE, JSPE, IEEJ, SICE and the IEEE computer society.

(Received May 27, 2020)

(Revised February 25, 2021)

Weakly Supervised Logo Detection Using a Dual-Attention Dilated Residual Network

Rahul Kumar JAIN[†](*Member*), Yutaro IWAMOTO[†], Taro WATASUE^{††}, Tomohiro NAKAGAWA^{††}, Takahiro SATO^{††}, Xiang RUAN^{††}, Yen-Wei CHEN[†]

[†] College of Information Science and Engineering, Ritsumeikan University, ^{††} tiwaki Co. Ltd., Kusatsu

<Summary> Automatic logo detection is a key function for many applications. Most existing logo detection methods are based on strong object level bounding box annotations, annotated either manually or automatically in synthesized images. However, in real applications, there are thousands of different types of logos with clutter background and varying sizes in images, that makes supervised detection methods less applicable because it is a labor-intensive task to do bounding box annotations. On the other hand, some weakly supervised learning-based methods have been proposed, their experimental results are not promising. In this paper, we propose weakly supervised methods for logo detection based on a dual attention dilated residual network (DRN) containing spatial and channel attention mechanisms using image-level labels instead of bounding box annotation for training data. The incorporated spatial attention module computes attention weights which are useful in predicting the spatial location in an image. Channel wise feature maps are used to emphasize the dependency of the channel in the global feature map and help to classify the logos into different categories. The use of attention-based mechanisms with DRN improves classification accuracy (around 4%) and considerably increases localization accuracy by more than 4% over a regular dilated residual network architecture.

Keywords: deep learning, weakly supervised localization, channel attention, spatial attention, dilated residual network

1. Introduction

Logo detection, which localizes and identifies logos in images, has become a very important functionality in many applications such as brand promotion, social media monitoring, auto-driving, intelligent transportation, market research and illegal logo detection, to name some of many.

However, logo detection is a challenging and difficult task in a complex scenario as there are many existing brand and logo images with diverse context. Logo is a proportionately small entity to detect in a large unconstrained image. Logos may contain unknown fonts (stylizes), different colors and logos can appear in varying sizes. Logos may also have inter-class similarity and intra-class differences, which makes this problem more difficult¹⁾.

Nowadays deep learning is being used in logo recognition because of superior performance in object detection. Deep learning requires a large training data with object level annotations to train a detector-model, which is not

only time-consuming and labor-intensive but also not affordable in some cases. To address this problem, in some approaches, the models have been trained with dataset consisting of synthesized training logo images to avoid manual annotation work^{2),3)}. But the performance was not sufficient when these models were tested on real images. On the other hand, the small number of logo images and classes in logo datasets affects the scalability of the model³⁾.

From the application perspective, learning from image level annotated data is more applicable and scalable than the object level annotated data. To address the bounding box annotation problem, in this paper, we propose a weakly supervised method using attention mechanisms based deep learning framework.

2. Related Work

Most of the recently proposed deep learning based logo recognition methods define the problem as a sub-task of general object detection, these methods use the datasets created manually or synthetically with object level anno-

tations for training. István et al.¹⁾ trained a class-agnostic detector⁴⁾ using logo and non-logo/background classes to identify logo object proposals. To retrieve logo images, they trained a separate network⁵⁾ by triplet loss using proxies⁶⁾. Su et al.²⁾ described the effect of the synthesized training data to deal with training data issue. In Ref.3), Su et al. proposed to use GANs methods for synthesizing coherent training logo images. Model self-learning principle using webly collected data has been implemented by Su et al.⁷⁾. They automatically collected most compatible training images from a webly noisy created data and trained a model iteratively by updating training dataset. Authors introduced model co-learning with last method in⁸⁾. They used two deep learning models^{4),9)} to identify the compatible training images. These identified training images have been fed as an input in cross-model manner.

Although in some methods, training images are synthetically or automatically created, the results do not show satisfactory performance on real images. A possible reason is that synthetically generated training images are not “real” enough to simulate their real-world counterparts, whereas automatically generated training data is very noisy. These methods are not much scalable. In this scenario we need more scalable deep learning based methods to address the issue of training the model with insufficient data. In real world, there are large number of existing logos. To train a detector we must annotate the logo images due to the data scarcity problem. To avoid such issues, weakly supervised training schemes can be used to build a more scalable model. In addition, attention mechanisms are very useful to locate and identify target regions in a weakly supervised manner.

Meanwhile spatial and channel attention mechanisms have been proposed and applied for various tasks such as image classification, object localization. Wang et al.¹⁰⁾ proposed a mask module implemented with a trunk branch that embeds bottom-up top-down feedforward structure to enhance spatial characteristics. In Ref.5), Hu et al. proposed to calculate channels statistical weights to enhance channel-wise feature responses. Woo et al.¹¹⁾ proposed to use a sequential combination of channel and spatial attention mechanisms. Channel and pixel-wise weights are determined using max-pooling and average-pooling operations. Chen et al.¹²⁾ incorporated a spatial and channel dual-attention mechanism for classification and localization of liver lesions on CT images. Among

many existing algorithms^{5),10),11),13)–15)}, our weakly supervised logo detection algorithm is based on a dual-attention dilated residual network (DADRN), which was proposed in our previous work¹²⁾. In our previous study, we have seen that DADRN has proved to be a more efficient feature extractor for weakly supervised localization than other networks. This network produces a large spatial resolution feature map which is more suitable for modeling small logo objects and helps to identify them. The proposed spatial and channel-wise attention mechanisms to generate and refine the robust final output feature map to activate the most probable candidate logo regions.

This approach can be summarized as: (I) We propose a weakly supervised logo detection algorithm which is capable to learn logo recognition with multiple logo classes without bounding box annotated training data. To our best knowledge this is the first weakly supervised training scheme for logo detection. We neither used bounding box annotations nor synthesized data for training, hence it is highly scalable to large scale real-world applications. (II) Logo image classification and logo detection are improved by implementing dual-attention based mechanism with the DRN Network.

3. Method

3.1 Overview

We present dilated residual network (DRN)¹⁶⁾ based architecture with attention mechanisms as feature extractor to generate a high resolution feature map. We employ spatial and channel attention mechanisms in different combinations to refine the final output feature maps. For weakly supervised localization task, class activation maps are generated using the Gradient-weighted method (Grad-CAM)¹⁷⁾ to visualize the possible logo regions.

3.2 Dilated Residual Network (DRN)

The architecture of DRN¹⁶⁾ follows the residual skip connections learning scheme with dilated convolution in higher layers. Dilated convolution operations cover a larger area than normal convolution operations and increase the receptive fields in higher layers. The last block of DRN network produces a high resolution global feature map (28×28), which provides better definition of image and improves the extracted features. **Figure 1** shows dilated convolution operation and DRN network. As depicted in the figure, ResNet network uses normal convolution operation ($d = 1$) while DRN uses an in-

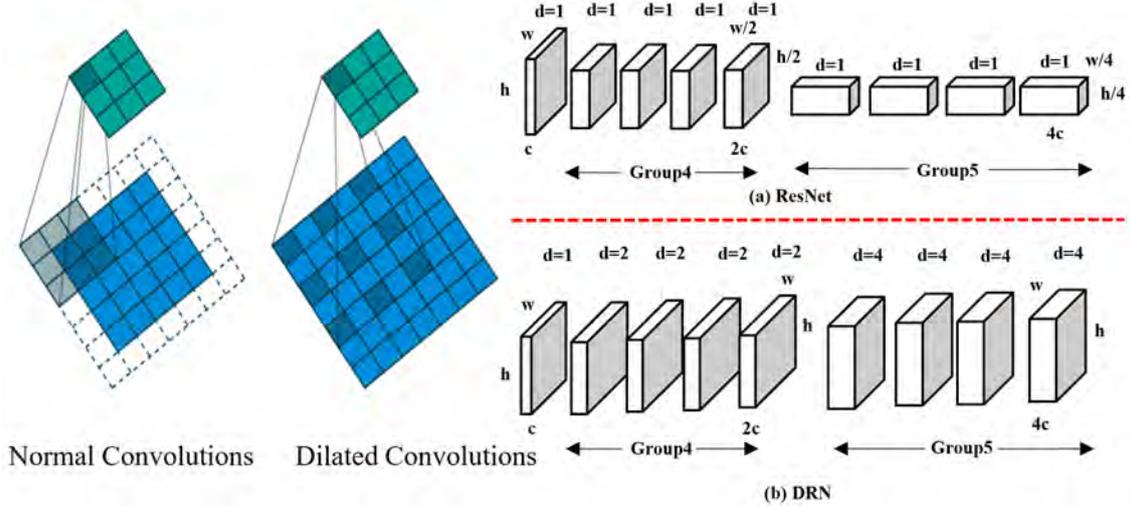


Fig. 1 DRN Network architecture

creasing factor of dilation ($d = 1, 2, 4$) in higher layers. The architecture of DRN and ResNet networks consists of five groups or blocks of convolutional layers. In Fig. 1, Group4 and Group5 represent the fourth and fifth convolutional blocks of the respective networks. Note that the DRN always shows better performance than ResNet in our previous work¹²⁾. So, we use DRN as our backbone network in this paper.

3.3 Spatial attention mechanism

To generate a spatial attention map, the global feature map $A \in R^{C \times H \times W}$ produced by DRN is fed into a 1×1 convolutional layer. The feature map is now transformed into two feature spaces B_1 and $B_2 \in R^{C1 \times (H \times W)}$. Where $C1$ shows the number of features and $(H \times W)$ is the size of flatten tensor. The value for each spatial position vector (P_j) is calculated using all other spatial position values in Eq.(1).

$$P_j = \sum_{i=1}^N \alpha_{j,i} D_i, \quad (1)$$

where $\alpha_{j,i}$ is the sigmoid normalization of $B_1^T B_2$. D_i is the transformed feature map after feeding into the 1×1 convolutional layer. $\alpha_{j,i}$ uses the importance of i^{th} position, while synthesising j^{th} position. The result of element-wise sum between the original feature map and attention feature map was the final output of the spatial attention mechanism, i.e.,

$$S_j = \beta P_j + A_j, \quad (2)$$

where β is a scale parameter, which is first initialized to 0 and assigned weight gradually during training process.

3.4 Channel attention mechanism

Channel based attention weights are calculated with respect to the global feature map. The values of attention

feature map (γ_{ji}) are the normalized values calculated using the sigmoid of the matrix multiplication between the feature map A and its transpose A^T . The channel attention map (C_j) is the element-wise sum of the original feature map and multiplication of the attention feature map with the original feature map.

$$C_j = \rho \sum_{i=1}^C (\gamma_{ji} A_i) + A_j. \quad (3)$$

Here $\gamma_{ji} = 1 / \{1 + e^{-(A \times A^T)}\}$, where $\gamma \in R^{C \times C}$ and ρ is a parameter which is first initialed with 0 and then gradually updated via back-propagation. The output feature maps of spatial (S_j) and chanel (C_j) attention modules are concatenated $[S_j, C_j]$ into $X \in R^{(C \times 2) \times H \times W}$ and passed as input to the classification branch. Where $C \times 2$ is the number of features of the concatenated attention maps and $(H \times W)$ is the spatial dimensions. Classification branch employs a global average pooling operation upon the concatenated attention feature maps. This branch has a 256-way fully-connected layer with Softmax to classify the logo images.

3.5 Gradient-weighted class activation map

The gradient-weighted class activation map (Grad-CAM) method¹⁷⁾ has been used to localize the logo regions. Grad-CAM generates the class activation map of the discriminative regions using the class specific prediction scores of channels in a feature map. In the Grad-CAM aactivation map, the gradient of the prediction score y^c is computed for class c with respect to the feature map activation $A^n (n \in [1, N])$, where N is the number of channels) of a convolutional layer, the gradient can be shown as $\partial y^c / \partial A_{ij}^n$. ReLU function is applied to generated weighted combination of activation maps to obtain the features with positive influence for any class c .

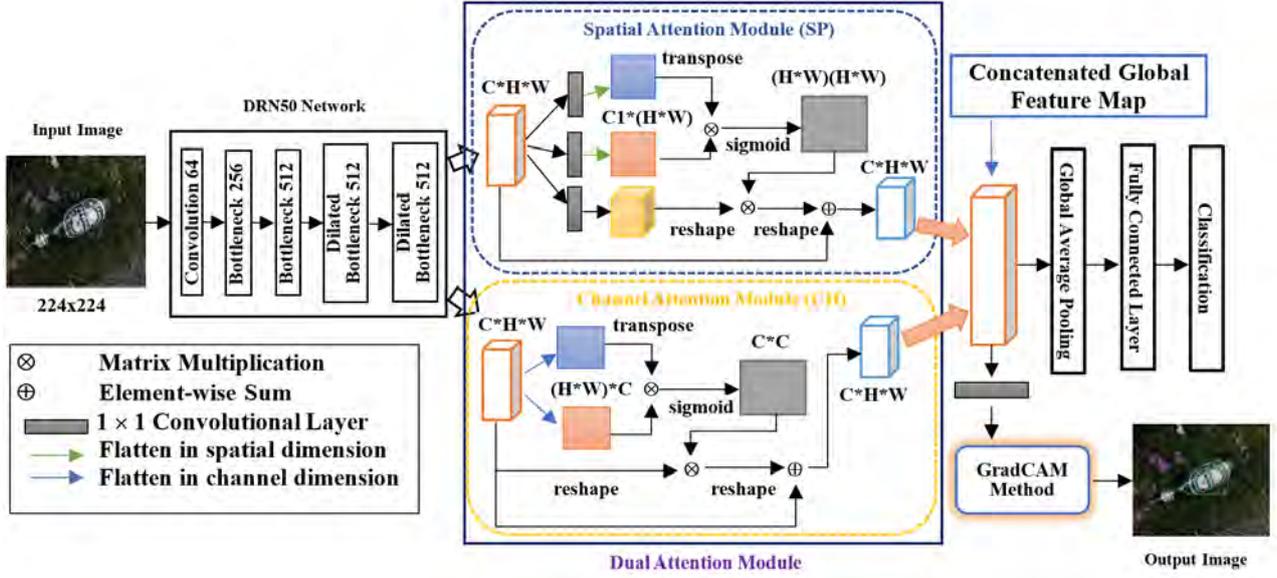


Fig. 2 DRN Network architecture with parallel attention mechanisms

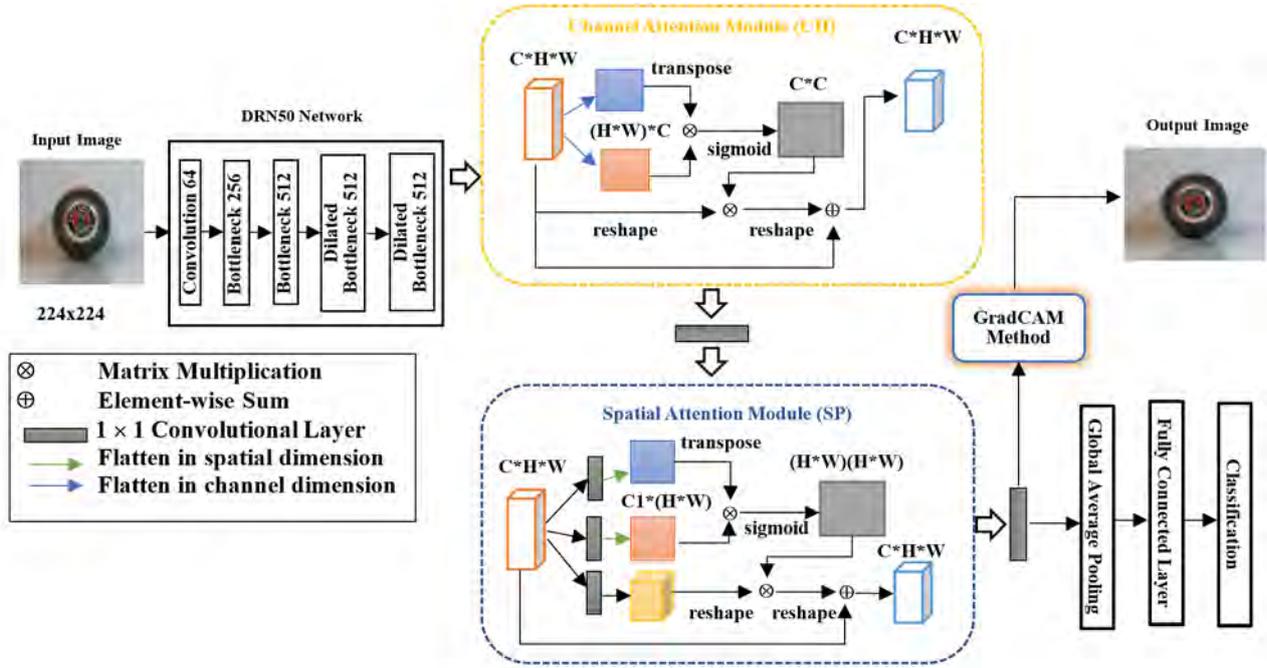


Fig. 3 DRN Network architecture with sequential attention mechanisms (CH=>SP)

$$L_{GradCAM}^c = ReLU \left(\sum_n \frac{1}{z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^n} A^n \right) \quad (4)$$

where $L_{GradCAM}^c$ is the generated activation map, z is the spatial dimension-size (height \times width) of feature maps.

We implement channel and spatial attention mechanisms in various combinations. We employ these modules separately, parallelly and sequentially along the network. Instead of only parallel combination of both attention modules as studied in previous work⁽¹²⁾, in this study we introduce sequential combinations of attention modules as an extension of the method. The main objective is

to investigate and explore the significance and applicability of the attention modules. **Figure 2** depicts the network architecture with attention mechanisms in parallel fashion, where both attention modules are implemented simultaneously and the attention weighted outputs are concatenated. **Figure 3** shows the sequential implementation of the channel (CH) and spatial (SP) attention modules with DRN network (i.e. CH => SP). The output of DRN feature extractor is first fed to the channel attention module and the refined output is sent through spatial attention module to generate the spatial attention maps.

4. Experiments and Results

4.1 Implementation

We implemented DRN50 framework in PyTorch with a mini-batch size of 16 and performed training for 200 epochs. For parameters optimization, we used Adam¹⁹) optimizer with an initial learning rate 0.0001 and a weight decay rate 0.1.

4.2 Dataset

We performed experiments on the logo images of FlickrLogos-32 dataset²⁰). FlickrLogos-32 dataset consists of 32 logo classes. There are 70 logo images in each class, where 40 images are for training-validation and 30 images are for test. In these 960 test images, there are total 1602 logo objects.

4.3 Evolution measures

We calculated classification accuracy for each network by determining the ratio of the number of total true positive cases to the total input test images in Eq.(5).

$$Classification_{acc} = \frac{\text{true positive predictions}}{\text{total predictions}} \quad (5)$$

For weakly supervised localization, a predicted bounding box was considered correct when Intersection over Union (IoU) exceeds 0.5 with the ground truth bounding box. We determined the localization accuracy using Eq.(6).

$$Localization_{acc} = \frac{\text{true positives}}{\text{total number of ground truths}} \quad (6)$$

4.4 Comparison with different attention modules

Table 1 shows classification accuracy for various networks: dilated residual network (DRN50), DRN50 network with parallel dual attention, DRN50 with separate spatial (SP) attention and channel (CH) attention module, DRN50 with sequential combinations of attention modules (i.e. channel and spatial attention module (CH=>SP), apatial and channel attention module (SP=>CH)), other existing methods SEResNet50⁵) and ResNet50 with CBAM¹¹).

The results show that DRN network performance is always better when attention modules are implemented in any combination. DRN50 network achieved around 84% accuracy rate. The classification accuracy of DRN network significantly increases by about 4% for parallel and sequential combinations of spatial and channel attention modules. We obtain more than 4% accuracy with spatial attention. While the accuracy rate improves by 2%

Table 1 Correctly classified logo images and classification accuracy

| Method | True Positives | Accuracy |
|---------------------------------|----------------|---------------|
| ResNet50 CBAM ¹⁵) | 783 | 81.563 |
| SEResNet ¹⁶) | 814 | 84.792 |
| DRN50 (w/o Attention) | 815 | 84.895 |
| DRN50 Channel Attention (CH) | 834 | 86.875 |
| DRN50 (Sequential: CH=>SP) | 846 | 88.125 |
| DRN50 (Sequential: SP=>CH) | 851 | 88.646 |
| DRN50 Dual Attention (Parallel) | 853 | 88.854 |
| DRN50 Spatial Attention (SP) | 857 | 89.271 |

Table 2 Correctly localized logo and localized accuracy

| Method | True Positives | Accuracy |
|---------------------------------|----------------|--------------|
| ResNet50 CBAM ¹⁵) | 241 | 15.04 |
| SEResNet ¹⁶) | 231 | 14.41 |
| DRN50 (w/o Attention) | 334 | 20.84 |
| DRN50 Spatial Attention (SP) | 389 | 24.28 |
| DRN50 Channel Attention (CH) | 388 | 24.21 |
| DRN50 (Sequential: CH=>SP) | 346 | 21.54 |
| DRN50 (Sequential: SP=>CH) | 311 | 19.41 |
| DRN50 Dual Attention (Parallel) | 401 | 25.03 |

using the channel attention mechanism. However, other methods achieve lower accuracy rate than our methods.

Table 2 depicts the mean localization accuracy and number of correctly located logo objects. This method focuses on a weakly-supervised logo localization training scheme based on attention mechanisms where the model is trained only with image-level annotations. Previously proposed methods in the field of logo detection use object level bounding box annotations for model training, so we did not find any weakly supervised based method for comparison.

Logo-object localization results are significantly better in comparison to the conventional DRN network. The accuracy increases with all combinations of the attention modules except for the sequential implementation of spatial and channel attention modules. We obtain more than 3% higher accuracy compared to conventional network when the attention modules are implemented separately. Parallel implementation of both attention module achieves the highest accuracy rate. It increases accuracy of the DRN network from 20.84% to 25.03%. While sequential implementation of channel and spatial attention mechanism enhances the accuracy by more than 1.5%.

The results indicate, the logo regions can be efficiently located using the dual-attention method. The parallel approach outperforms sequential and conventional approaches because: 1) parallel approach simultaneously refines the output feature maps in spatial and channel di-

mensions. 2) concatenation of feature maps provides the important attention weights to learn classification and localization tasks more accurately. Whereas, in sequential approaches, attention-weighted feature maps are refined independently one after the other, reducing the effectiveness of important weights during learning.

In the parallel implementation, the final output was a concatenated feature map of those generated by the channel and spatial attention module. Out of 256 channels, the first 128 channels were generated by the spatial attention module and the channel numbers 129 to 256 were generated by the channel attention module. We computed channel specific response of channel attention module for each category. The top 5 channels with higher weights for each category are given in **Table 3**. The channel specific response indicates the importance of channels with the corresponding category. After analyzing the table, we found that five channels with corresponding numbers [211, 231, 229, 213, 217] occur more frequently in all categories. Analysis of the occurrence of the channels suggests these channels play a major role in identifying

logo regions. This study can be useful to assign more biased weights to these specific channels to improve the performance.

Figure 4 visualizes the channel-specific response of top five channels (left to right - 211, 231, 229, 213, 217) and the final heatmaps, respectively. These binary images illustrate different attention-weights having different corresponding response. The final heatmaps are generated by calculating Grad-CAM visualization for the last convolutional outputs. Grad-CAM visualization uses the positive influence of all channels to locate the possible logo regions.

Figure 5 depicts the class activation regions of various methods generated through Grad-CAM^[7]. These methods are DRN50, DRN50 with parallel attention modules, DRN50 with spatial attention, ResNet50 with CBAM attention block. Our proposed dual-attention modules in parallel combination generates qualitative results with more precise target specific class activation and enhances object localization capability of feature extractor.

In practice, all bounding box-based object detectors face the major problem of lack of sufficient object-level annotated training data. Creating datasets with manually drawn object level bounding boxes is time and labor consuming. Training data with only image level annotations is more useful in this scenario. Although the performance of the logo localization is not compatible with methods that use object-level bounding boxes for training, our weakly supervised method is considerable and easily scalable to a large number of logo classes.

Table 3 Channel specific response for DRN50 Parallel Dual-Attention network

| | |
|--------------|---------------------------------|
| Adidas | 211, 231, 238, 251, 229, |
| Aldi | 231, 169, 197, 199, 129, |
| Apple | 231, 233, 238, 132, 251, |
| Becks | 229, 217, 211, 190, 227 |
| Bmw | 213, 149, 221, 231, 233 |
| Carlsberg | 146, 211, 147, 213, 245 |
| Chimay | 229, 147, 248, 216, 132 |
| Cocacola | 231, 129, 217, 190, 211 |
| Corona | 211, 229, 169, 245, 248 |
| Dhl | 213, 129, 197, 251, 221 |
| Erdinger | 211, 169, 251, 232, 213 |
| Esso | 217, 202, 211, 214, 233 |
| Fedex | 217, 213, 233, 241, 229 |
| Ferrari | 216, 231, 238, 215, 207 |
| Ford | 211, 245, 190, 216, 213 |
| Fosters | 211, 189, 221, 216, 213 |
| Goolge | 229, 231, 190, 241, 254 |
| Guinness | 211, 219, 146, 248, 245, |
| Heineken | 217, 146, 213, 238, 229 |
| HP | 231, 229, 240, 190, 197 |
| Milka | 211, 241, 233, 151, 231 |
| Nvidia | 229, 153, 238, 145, 146 |
| Paulaner | 169, 213, 221, 135, 236 |
| Pepsi | 213, 233, 217, 189, 251 |
| Rittersport | 243, 145, 229, 185, 238 |
| Shell | 231, 251, 197, 229, 129 |
| Singha | 232, 211, 189, 169, 238 |
| Starbucks | 238, 227, 221, 169, 189 |
| Stellaartois | 227, 211, 189, 129, 147 |
| Texaco | 217, 197, 227, 190, 151 |
| Tsingtao | 189, 241, 177, 217, 213 |
| Ups | 233, 211, 145, 251, 245 |

5. Conclusion

In this paper, we proposed a spatial and channel attention mechanism based network for logo classification and localization. This method is scalable to a large number of logo images with different classes and it needs image-level annotated data for training the model. The results show better classification and localization accuracy over conventional architecture. This method is highly scalable to deal with real-world problems.

References

- 1) I. Fehérvári, S. Appalaraju: "Scalable Logo Recognition Using Proxies", Proc. of IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 715–725 (2019).
- 2) H. Su, X. Zhu, S. Gong: "Deep Learning Logo Detection with Data Expansion by Synthesising Context", Proc. of IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 530–539 (2017).
- 3) H. Su, X. Zhu, S. Gong: "Open Logo Detection Challenge",

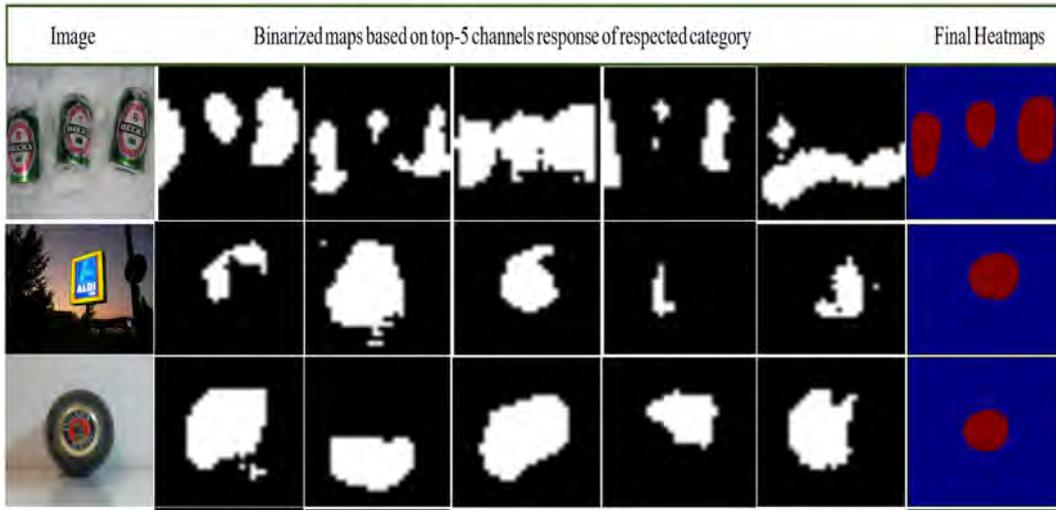


Fig. 4 Channel specific response for different logo objects

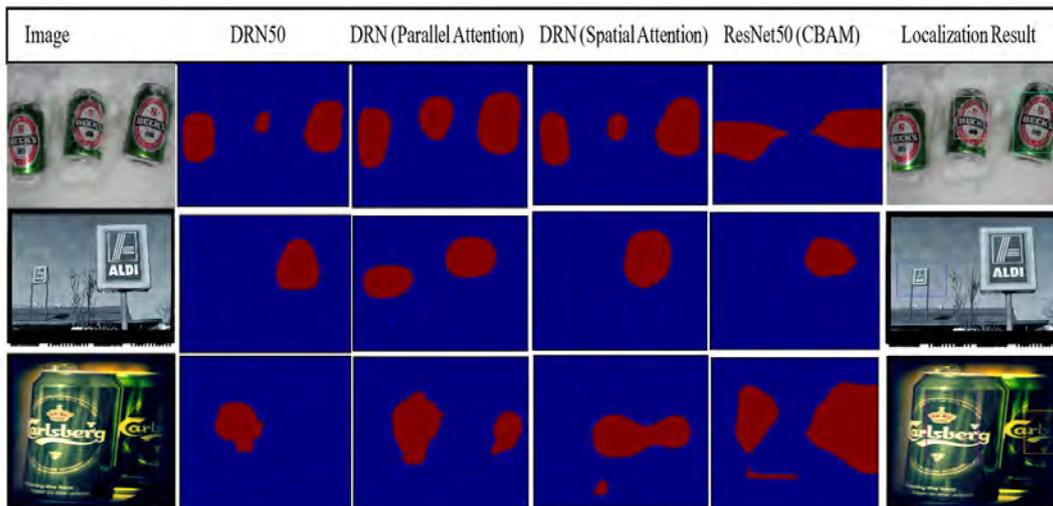


Fig. 5 Class activation map visualization using Grad-CAM

Proc. of British Machine Vision Conference (BMVC) (2018).

- 4) S. Ren, K. He, R. Girshick, J. Sun: "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 1137–1149 (2017).
- 5) J. Hu, L. Shen, G. Sun: "Squeeze and Excitation Networks", *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7132–7142 (2017).
- 6) Y. Movshovitz-Attias, A. Toshev, T. K. Leung, S. Ioffe, S. Singh: "No Fuss Distance Metric Learning Using Proxies", *Proc. of IEEE International Conference on Computer Vision (ICCV)*, pp. 360–368 (2017).
- 7) H. Su, X. Zhu, S. Gong: "Weblogo-2M: Scalable Logo Detection by Deep Learning from the Web", *Proc. of IEEE International Conference on Computer Vision Workshops (ICCVW)*, pp. 270–279 (2017).
- 8) H. Su, X. Zhu, S. Gong: "Scalable Logo Detection by Self Co-Learning", *Pattern Recognition*, Vol.97, 107003 (2020).
- 9) J. Redmon, A. Farhadi: "Yolo9000: Better, Faster, Stronger", *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6517–6525 (2017).
- 10) F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, X. Tang: "Residual Attention Network for Image Classification", *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6450–6458 (2017).
- 11) S. Woo, J. Park, J. Y. Lee, I. S. Kweon: "CBAM: Convolutional Block Attention Module", *Proc. of European Conference on Computer Vision (ECCV)*, pp. 3–19 (2018).
- 12) X. Chen, L. Lin, D. Liang, H. Hu, Q. Zhang, Y. Iwamoto, X.H. Han, Y.W. Chen, R. Tong, J. Wu: "A Dual-Attention Dilated Residual Network for Liver Lesion Classification and Localization on CT Images", *Proc. of IEEE International Conference on Image Processing (ICIP)*, pp. 235–239 (2019).
- 13) J. Choe, H. Shim: "Attention-Based Dropout Layer for Weakly Supervised Object Localization", *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.2219–2228 (2019).
- 14) K. Yang, D. Li, Y. Dou: "Towards Precise End-to-end Weakly Supervised Object Detection Network", *Proc. of IEEE International Conference on Computer Vision (ICCV)*, pp. 8372–8381 (2019).
- 15) H. Xue, C. Liu, F. Wan, J. Jiao, X. Ji, Q. Ye: "DANet: Divergent Activation for Weakly Supervised Object Localization", *Proc. of IEEE International Conference on Computer Vision (ICCV)*, pp. 6589–6598 (2019).
- 16) F. Yu, V. Koltun, T. Funkhouser: "Dilated Residual Networks", *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 472–480 (2017).

- 17) R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra: "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization", Proc. of IEEE International Conference on Computer Vision (ICCV), pp. 618–626 (2017).
- 18) K. He, X. Zhang, S. Ren, J. Sun: "Deep Residual Learning for Image Recognition", Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 (2016).
- 19) Y. Bengio, Y. LeCun: "Adam: A Method for Stochastic Optimization", Proc. of 3rd International Conference on Learning Representations (ICLR) (2015)
- 20) S. Romberg, L. G. Pueyo, R. Lienhart, R. Van Zwol: "Scalable Logo Recognition in Real-World Images", Proc. of the 1st ACM International Conference on Multimedia Retrieval, pp. 1–8 (2011).

(Received June 12, 2020)
 (Revised January 6, 2021)



Rahul Kumar JAIN (*Member*)

He received BCA degree from Dr. H. S. Gour Central University, Sagar, India, and MCA degree from RGPV University, Bhopal, India in 2011 and 2015, respectively. He is now pursuing PhD at Ritsumeikan University, Kusatsu, Japan.



Yutaro IWAMOTO

He received the B.E. and M.E., and D.E. degrees from Ritsumeikan University, Kusatsu, Japan in 2011 and 2013, and 2017, respectively. He is currently an Assistant Professor at Ritsumeikan University, Kusatsu, Japan. His current research interests include image restoration, segmentation, classification of medical imaging and deep learning.



Taro WATASUE

He received the B.E. and M.E. degrees from Osaka University, Osaka, Japan, in 1998 and 2002, respectively. From 2004 to 2007, he was an assistant professor in Kobe University. From 2007 to 2016, he was an Expert Engineer with TOME R&C Inc., Kyoto, Japan. He is currently the co-founder of tiwaki Co., Ltd., Kusatsu, Japan. His current research interests include computer vision, machine learning, and image processing.



Tomohiro NAKAGAWA

He has experience in R & D of machine learning and computer vision since 1998. He is currently the co-founder of tiwaki Co., Ltd., Kusatsu, Japan. Since 2019, he has been involved in developing pose estimation technology that runs over 50fps on Jetson nano.



Takahiro SATO

He received the B.E., M.E, and Ph.D. degrees from Hiroshima University, Hiroshima, Japan in 2006, 2008 and 2011, respectively. From 2011 to 2016, he was an Expert Engineer with TOME R & C Inc., Kyoto, Japan. He is currently an Expert Engineer with tiwaki Co., Ltd., Kusatsu, Shiga, Japan. His current research interests include computer vision, machine learning, and image processing.



Xiang RUAN

He received the B.E. degree from Shanghai Jiao Tong University, Shanghai, China, in 1997, and the M.E and Ph.D. degrees from Osaka City University, Osaka, Japan, in 2001 and 2004, respectively. From 2007 to 2016, he was an Expert Engineer with OMRON Corporation, Kyoto, Japan. He is currently the Founder and CEO of tiwaki Co., Ltd., Kusatsu, Japan. His current research interests include computer vision, machine learning, and image processing.



Yen-Wei CHEN

He received his B.E. degree in 1985 from Kobe University, Kobe, Japan. He received his M.E. degree in 1987, and his D.E. degree in 1990, both from Osaka University, Osaka, Japan. He is currently a professor at the college of Information Science and Engineering, Ritsumeikan University, Japan. His research interests include computer vision, medical image analysis.

Interactive Control of Fire Color based on Physicochemical Combustion

Ryosuke UCHIKAWA[†] (*Student Member*), Hiroki ISHIZUKA[†], Sei IKEDA[†], Yoshihiro KURODA^{††}, Shunsuke YOSHIMOTO^{†††}, Osamu OSHIRO[†]

[†]Osaka University, ^{††}University of Tsukuba, ^{†††}The University of Tokyo

<Summary> The expression of flames using computer graphics is frequently required as a familiar phenomenon. Various colors, including red and blue, can be seen in real flames. However, the combustion reaction in a real flame is too complex, involving thousands of chemical reactions to render interactive simulations. In this study, the authors constructed a novel combustion reaction model for expressing realistic flames with simplified chemical reactions of complex combustion. To achieve interactive physicochemical simulations of flames, fluid physics simulations were applied using the position-based fluids method. In the proposed combustion model, the combustion reaction is represented as five dominant chemical reactions. Opposed to the conventional methods, the proposed model can reproduce the blue and red regions of the flame. The proposed flame simulation method can be operated at approximately 20 frames per second or more, making interactive rendering possible.

Keywords: flame, combustion, interactive, position-based fluids (PBF), position-based dynamics (PBD), computer graphics (CG)

1. Introduction

In most conventional computer graphics (CG) methods expressing flames, flames are represented only by the colors around red and yellow spectral bands derived from the black-body radiation of the soot produced by the combustion reaction^{1)–5)}. However, real flames apparently contain colors on different bands, such as blue. For example, the flames from gas cooking burners are turquoise blue, the flame of burning alcohol is pale blue, and the base of a lighter flame is blue, whereas its tip is yellow or red. Blue flames originate as the luminescence of intermediate components produced when the fuel decomposes and luminescence of carbon monoxide combustion. Nguyen et al.⁶⁾ proposed a simulation method that considers the blue region existing at the base of a flame. However, this method only allows a single intensity of the flame's blue color; thus, it is expressed by only a single color.

The luminescence intensity of the generated chemical species must be calculated based on the computation of the combustion reaction to reproduce a more plausible color of the flame. However, the combustion reaction consists of thousands of chemical sub-reactions and cannot be accurately reproduced in real-time simulation. Ihm et al.⁷⁾ and Kang et al.⁸⁾ constructed a combustion reac-

tion model and simulated a flame based on the hydrogen combustion mechanism, which can be represented by five chemical reaction equations. However, this model does not consider the generation of the chemical species that are important in determining the flame's color or reproduce the flame's blue color.

In this study, we aim to establish an interactive method for simulating a realistic flame having various colors, including blue and red, originating from the luminescence. To achieve this, we propose a novel physicochemical combustion reaction model suitable for reproducing the flame's color. We also propose a method for interactive physicochemical simulation of flames. For this simulation, we combined the proposed combustion reaction model and the position-based fluids (PBF) model⁹⁾, which is a high-speed fluid calculation method.

Our simulation of the combustion reaction can be divided into two stages: decomposition and synthesis. In the decomposition stage, fuel molecules are decomposed to the atomic level in one chemical reaction. In the synthesis stage, the final products of combustion, that is, carbon monoxide, carbon dioxide, and soot, are determined by referring to the concentration and temperature of each species. Using this method, the calculation time for obtaining the information necessary for reproducing

the flame color, such as the fuel decomposition rate and soot concentration can be shortened. The PBF model is used to calculate the air flow. The fluid simulation was performed with the constraint of constant density on each particle, rather than calculating the pressure term of the Navier–Stokes equation. This method increases the simulation time step width, decreases the computation time, and enables interactive calculation. Furthermore, to reproduce the flames produced by the combustion of gaseous and solid fuels, such as matchsticks, the solid is expressed as a mass of particles, and the thermal decomposition reaction inside the solid particles is considered.

An interactive flame simulation method is presented in section 3 and section 4. In particular, a method for achieving a physicochemically accurate reproduction of flame colors is described. The representation of solid fuel combustion is also explained. After the implementation of the proposed model is shown in section 4, the examples of solide fuel combustion are shown in section 5.

2. Related Work

The red and yellow colors of the flame are due to soot blackbody radiation, and the blue of the flame is due to light emission when the fuel decomposes or when carbon monoxide burns¹⁰⁾. If the combustion reaction simulation can reproduce this phenomenon, it is possible to draw a plausible flame. However, the combustion reaction is a complex reaction in which several thousand elementary reactions occur serially. Many researchers have attempted to model combustion reactions. One of the most interesting methods is the flame front model, which handles the combustion reaction by assuming that the combustion reaction rate is sufficiently fast compared to the diffusion of chemical species and heat, so that the combustion reaction occurs only on a thin surface (flame front)¹¹⁾. Nguyen et al.⁶⁾ combined a flame front model with a level set method to produce a realistic flame without considering the details of the combustion reaction. In addition, Hong et al.¹²⁾ applied explosion theory to the method of Nguen et al.⁶⁾ to allow the representation of wrinkled flames by using simulations. However, in these flame front models, although the location of the combustion reaction can be specified, the reaction speed and the type of reaction cannot be considered, and thus, the color of the flame cannot be accurately reproduced. A realistic flame can be interactively generated if a new combustion reaction model could reproduce the color of the flame at high speed.

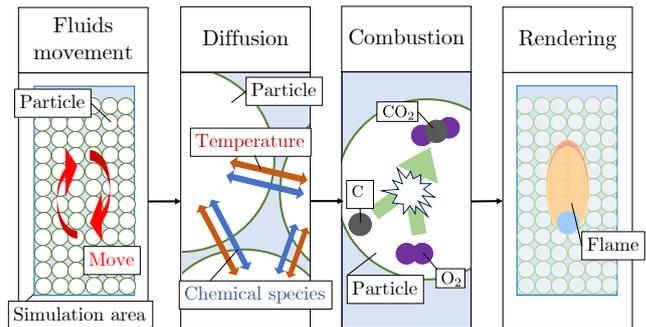


Fig. 1 Overview of the proposed simulation method

Studies in which the combustion reaction was modeled directly, and the flame was simulated accordingly have been conducted. Min et al.¹³⁾ and Stam et al.¹⁴⁾ simulated flames by modeling the complex combustion reactions using one chemical reaction equation. However, because this method treats only one chemical reaction, it cannot effectively include flame characteristics, such as temperature and soot concentration distributions, and the resulting CG reproduction cannot be considered realistic. Ihm et al.⁷⁾ and Kang et al.⁸⁾ modeled combustion based on the simplest hydrogen combustion reaction mechanism, which consists of five chain reactions. However, even in this model, factors such as soot concentration and the burning rate of carbon monoxide, which are essential for rendering a flame in CG, were not addressed.

3. Method

To render a flame using simulation, it is necessary to calculate the flow of the atmosphere around the flame and the combustion reaction inside the flame. We simulate the flames by discretizing air with particles and tracking their behavior. Our system consists of four parts, as shown in **Fig. 1**. The first part is a numerical fluid calculation unit that calculates the gas flow. The second is the diffusion calculation unit that calculates the diffusion of each species and temperature, The third is a combustion reaction calculation unit that calculates the combustion reaction in each particle, and the fourth is a rendering unit for expressing a flame using the numerical values obtained by calculation. In the following sections, we first describe the calculation of fluid motion and diffusion. Then, we discuss the modeling and calculation of the combustion reaction. Finally, the flame rendering method is explained.

3.1 Fluid motion computation

In this section, we describe an interactive fluid motion calculation method. We first describe position-based dynamics (PBD), which is a fast and stable physical simulation method, and then PBF, which applies PBD to fluid motion computation. Finally, we discuss the treatment of buoyancy in the PBF method.

3.1.1 Position-based dynamics algorithm

PBD is a method of obtaining the position of a calculation point directly using a constraint condition. Let \mathbf{x}_i , \mathbf{v}_i , m_i and \mathbf{f}_i be the the position, velocity, mass and external force of the i -th particle. The total number of particles is n . \mathbf{x} is $(\mathbf{x}_0, \dots, \mathbf{x}_{n-1})$, \mathbf{v} is $(\mathbf{v}_0, \dots, \mathbf{v}_{n-1})$, and \mathbf{f} is $(\mathbf{f}_0, \dots, \mathbf{f}_{n-1})$.

The process at each step in the basic PBD is as follows. The predicted velocity \mathbf{v}^* is determined as $\mathbf{v}^* = \mathbf{v} + \mathbf{f}/\mathbf{m}\Delta t$, where Δt is the time step, $\mathbf{f}/\mathbf{m} = (\mathbf{f}_0/m_0, \dots, \mathbf{f}_{n-1}/m_{n-1})$, and the predicted position \mathbf{x}^* is determined by $\mathbf{x}^* = \mathbf{x} + \mathbf{v}^*\Delta t$. Then, to find the corrected position $\tilde{\mathbf{x}}$ that satisfies the constraint condition $C(\cdot)$, as explained below, $\Delta\mathbf{x} = \tilde{\mathbf{x}} - \mathbf{x}^*$ is obtained by solving

$$C(\mathbf{x}^* + \Delta\mathbf{x}) = C(\mathbf{x}^*) + \sum_{k < n} \nabla_{\mathbf{x}_k} C(\mathbf{x}^*) \cdot \Delta\mathbf{x}_k = 0, \quad (1)$$

where $\nabla_{\mathbf{x}_k} = (0, \dots, \partial/\partial x_k, \partial/\partial y_k, \partial/\partial z_k, \dots, 0)$. To satisfy the conservation of momentum and angular momentum, $\Delta\mathbf{x}_k$ is expressed as

$$\Delta\mathbf{x}_k = \lambda m_k^{-1} \nabla_{\mathbf{x}_k} C(\mathbf{x}^*), \quad (2)$$

where λ is a variable. From Eqs. (1) and (2), λ can be obtained as follows:

$$\lambda = - \frac{C(\mathbf{x}^*)}{\sum_{k < n} m_k^{-1} |\nabla_{\mathbf{x}_k} C(\mathbf{x}^*)|^2}. \quad (3)$$

Finally, the particle velocity from the corrected position $\tilde{\mathbf{x}}$ and initial position \mathbf{x} is updated. In general, PBD methods use Jacobi or Gauss-Seidel iterations to converge the constraints.

3.1.2 Position-based fluids

In the PBF method, fluid simulation is achieved by imposing density constraints on all particles¹⁵⁾. The density constraint equation for particle i in the PBF method is expressed as

$$C_i(\mathbf{x}) = \frac{\rho_i}{\rho_i^0} - 1 \leq 0, \quad (4)$$

where ρ_i^0 is the rest density of the target fluid, and ρ_i is the density of particle i , calculated as

$$\rho_i = \sum_{j < n} m_j W_{ij}. \quad (5)$$

$W_{ij} = W(\mathbf{x}_i - \mathbf{x}_j, \xi)$ is the kernel function, where ξ is the effective particle radius. In the original PBF⁹⁾, density constraint $C_i(\mathbf{x}) = 0$ was used, but a non-negative clamping density constraint is sufficient for representing fluid motion¹⁵⁾. From Eqs. (1) to (5), the position correction vector $\Delta\mathbf{x}_i$ of particle i can be obtained as

$$\Delta\mathbf{x}_i = \sum_{j < n} \left(\frac{\lambda_i}{m_i \rho_i^0} + \frac{\lambda_j}{m_j \rho_j^0} \right) m_j \nabla_{\mathbf{x}_i} W_{ij}. \quad (6)$$

3.1.3 Buoyancy

The following two factors are considered as the main factors of buoyancy associated with gas combustion.

- Nonuniform density distribution due to fuel supply and chemical species diffusion.
- Nonuniform temperature distribution due to combustion reaction and temperature diffusion.

In the PBF method, the buoyancy generated between fluids with different densities is expressed by the difference in the rest density ρ^0 and the particle mass m_i ¹⁵⁾. Therefore, we update m_i and ρ_i^0 for each step:

$$m_i = \sum_{\ell < \Theta} \mathbf{D}_\ell d_{i,\ell}, \quad (7)$$

$$\rho_i^0 = \frac{\sum_{\ell < \Theta} \mathbf{D}_\ell d_{i,\ell}}{\sum_{\ell < \Theta} d_{i,\ell}} \frac{p_0}{R_0 T_0}, \quad (8)$$

where $d_{i,\ell}$ is the molar amount of the chemical species ℓ in particle i . Θ is the total number of chemical species used in the simulation, \mathbf{D}_ℓ is the mass per mole of the chemical species ℓ , p_0 is the reference pressure, R_0 is the gas constant, and T_0 is the initial temperature.

The buoyancy due to the temperature distribution should be expressed as the force generated by the density change; however, if a large difference exists in the density ρ^0 of the particles, the calculation diverges and the value becomes extremely large. Therefore, we implemented buoyancy due to temperature, not as density change, but as an external force acting on the particles⁶⁾. Let T_i^{amb} be the temperature around particle i . Then, we express the buoyancy $\mathbf{F}_i^{\text{buo}}$ applied to particle i as

$$\mathbf{F}_i^{\text{buo}} = \vartheta (T_i - T_i^{\text{amb}}) \mathbf{g}, \quad (9)$$

where ϑ is the buoyancy constant. The ambient temperature T_i^{amb} is given by

$$T_i^{\text{amb}} = \frac{\sum_{j \neq i} T_j W_{ij}}{\sum_{j \neq i} W_{ij}}. \quad (10)$$

3.1.4 Viscosity

The contribution of viscosity to the shape of the flame is considered non-negligible. We applied the artificial method¹⁶⁾ used in SPH to consider changes in motion due to viscosity. Although this is not the physical viscosity, it is enough to represent the viscosity effect and is easier to tune.

$$\mathbf{v}_i^{\text{new}} = \mathbf{v}_i + \zeta \sum_{j < n} (\mathbf{v}_j - \mathbf{v}_i) W_{ij}, \quad (11)$$

where $\mathbf{v}_i^{\text{new}}$ is the new velocity of particle i , and ζ is the tunable parameter set as $\zeta = 0.0002$ in our simulations.

3.2 Diffusion

This section describes the diffusion of the chemical species and temperature. The diffusion equation of a generalized physical quantity ϕ can be expressed as

$$\frac{\partial \phi}{\partial t} = \nabla \cdot (D \nabla \phi), \quad (12)$$

where D is the diffusion coefficient of ϕ , and t is the time. In our simulation, ϕ is the temperature T or amount of chemical species d .

For gas concentration diffusion and temperature diffusion, it is assumed that the diffusion coefficient D can be expressed as $D = D_0 T / T_0$ using the diffusion coefficient D_0 at the reference temperature. Equation (12) then becomes

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \nabla D \cdot \nabla \phi + D \nabla^2 \phi \\ &= D_0 \frac{1}{T_0} \nabla T \cdot \nabla \phi + D_0 \frac{T}{T_0} \nabla^2 \phi. \end{aligned} \quad (13)$$

In our simulation, $\nabla T \approx \nabla \phi$ and $1 \ll T$, we can assume that $D_0 / T_0 \nabla T \ll D_0 T / T_0 \nabla \phi$ and ignore the first term of Eq.(13) and Eq.(14) is derived.

$$\frac{\partial \phi}{\partial t} = D_0 \frac{T}{T_0} \nabla^2 \phi \quad (14)$$

Assuming that the temperature diffusion and concentration diffusion are caused by the gradient to maintain symmetry, Eq.(15) is obtained by discretizing Eq.(14) using the kernel function and forward Eulerian form.

$$\phi_i^{\text{new}} = \phi_i + \Delta t \sum_{j < n} m_j D_0 \frac{T_i + T_j}{2T_0} \frac{\phi_j - \phi_i}{\rho_j} \nabla^2 W_{ij}. \quad (15)$$

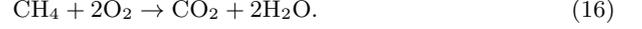
We use Eq.(15) to handle the diffusion phenomena of the gas temperature and concentration of each chemical species.

3.3 Combustion reaction modeling

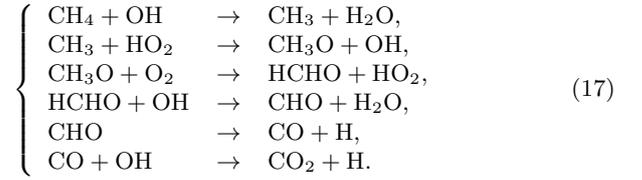
This section describes the combustion reaction and proposes a novel combustion reaction model for rendering

each flame. Moreover, we describe the application of the proposed combustion reaction model to solid fuel combustion.

The reaction in which methane gas burns to form carbon dioxide is represented as



However, this represents only the beginning and the end of the combustion reaction, and does not represent the reaction process. The combustion reaction occurring in an actual flame is expressed as a set of reactions, for example:



Formula (17) shows one of the pathways by which one methane molecule is oxidized to carbon dioxide. The specific processes that occur in each chemical species that make up Fm.(17) are called elementary reactions.

Combustion is a complex chemical phenomenon involving many elementary reactions; approximately 400 elementary reactions are involved even in the case of low-carbon lower alcohols, such as ethanol¹⁷⁾. The number of elementary reactions increases exponentially as the carbon number increases. Thus, it is difficult to simulate all the elementary reactions involved in the combustion of each particle. However, it is known that the blue color of a flame is due to the light emission of the chemical species generated when the fuel is decomposed and the light emission when carbon monoxide is burned to form carbon dioxide. Similarly, the red color of a flame is due to the black body radiation of the soot resulting from the combustion reaction. Therefore, the important information required to render a flame is the soot concentration, decomposition rate of the fuel, and burning rate of carbon monoxide. To obtain this information without considering the entire reaction, we consider a simplified model of the combustion reaction, as shown in **Fig. 2**. In the next subsections, we explain our combustion model, which is composed of these chemical reactions.

3.3.1 Fuel decomposition

Figure 2(a) shows the decomposition of gaseous fuel. The decomposition reaction is generated by the collision of fuel molecules in high heat with other molecules; it is assumed that the fuel molecules are decomposed to the

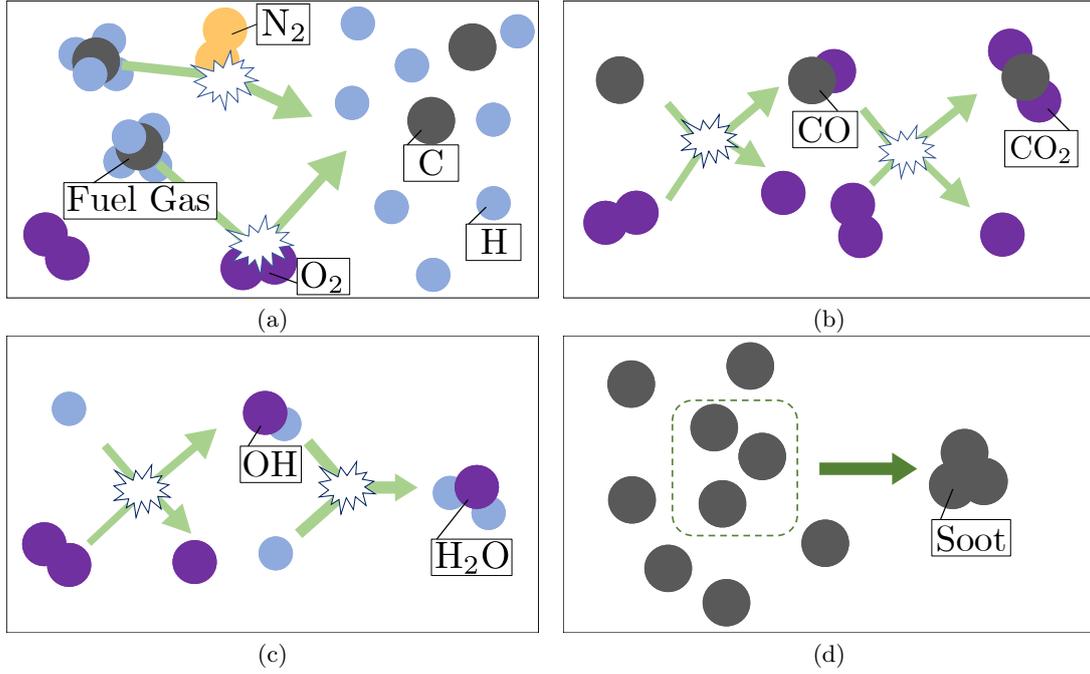
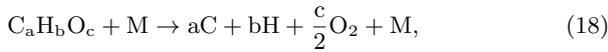


Fig. 2 Proposed combustion reaction model; (a)Decomposition of gaseous fuel, (b)Combustion mechanism of the generated carbon atom, (c)Process of the generated hydrogen atoms to become water, and (d)Soot generation model

atomic level in one collision. Assuming that the fuel is represented as $C_aH_bO_c$, the reaction formula is



where M denotes a third body. To simplify the decomposition of fuel and the following chemical reactions, we consider the generated oxygen atoms become oxygen molecules immediately and do not react with other molecules. We consider that this simplified model is more complicated than the conventional models and can reproduce complex phenomena. Assuming that Eq.(18) is a second-order reaction, the reaction rate u_{fuel} is determined as

$$u_{fuel} = \kappa_{fuel}[fuel][M], \quad (19)$$

where $[fuel]$ is the concentrations of the fuel. κ_{fuel} is the reaction rate constant, obtained by the Arrhenius Eq.(20).

$$\kappa = A \exp\left(-\frac{E_a}{R_0T}\right), \quad (20)$$

where A is a frequency factor and E_a is the activation energy, which varies depending on the reaction equation of interest.

3.3.2 Carbon combustion

Figure 2(b) shows the combustion mechanism of the generated carbon atom. We model the mechanism by

which carbon monoxide and carbon dioxide are generated from carbon atoms and oxygen molecules as follows



We assume that the generated oxygen atoms become oxygen molecules immediately. Both Fms.(21) and (22) represent secondary reactions. The reaction rates u_{CO} and u_{CO_2} were determined by

$$u_{CO} = \kappa_{CO}[C][O_2], \quad (23)$$

$$u_{CO_2} = \kappa_{CO_2}[CO][O_2]. \quad (24)$$

3.3.3 Hydrogen combustion

Figure 2(c) shows the process by which the generated hydrogen atoms become water, which is one of the final products. It is assumed that the mechanism of water formation from hydrogen atoms and oxygen molecules consists of two elementary reactions:



Furthermore, it is assumed that the reaction in Eq.(26) is sufficiently fast compared to the reaction in Eq.(25) and that the generated OH is consumed instantaneously. By applying the steady-state approximation to OH , we can determine a series of reaction rates u_{H_2O} as

$$u_{H_2O} = \kappa_{H_2O}[H][O_2]. \quad (27)$$

3.3.4 Soot generation

Figure 2(d) shows the soot generation model. In our combustion model, soot is formed by the concentration of carbon that does not react with oxygen. The formation process of this soot is not described by a reaction equation, and it is assumed that the soot concentration is proportional to the carbon concentration.

3.4 Chemical species concentration

We handle combustion reactions by dividing them into decomposition and synthesis. Therefore, the chemical species concentration is updated in two steps. First, the rate of the decomposition reaction shown in Eq.(18) is calculated inside each particle, and the amount of each chemical species is updated:

$$\frac{d[\text{fuel}]}{dt} = -u_{\text{fuel}}, \quad (28)$$

$$\frac{d[\text{C}]}{dt} = au_{\text{fuel}}, \quad (29)$$

$$\frac{d[\text{H}]}{dt} = bu_{\text{fuel}}, \quad (30)$$

$$\frac{d[\text{O}_2]}{dt} = \frac{c}{2}u_{\text{fuel}}. \quad (31)$$

Next, after updating the concentration of each species using Eq.(28) to Eq.(31), we obtain the rate of the synthetic reaction shown in Eqs.(21), (22) and (25), and the concentration of the chemical species is next calculated:

$$\frac{d[\text{C}]}{dt} = -u_{\text{CO}}, \quad (32)$$

$$\frac{d[\text{H}]}{dt} = -2u_{\text{H}_2\text{O}}, \quad (33)$$

$$\frac{d[\text{O}_2]}{dt} = -0.5(u_{\text{CO}} + u_{\text{CO}_2} + u_{\text{H}_2\text{O}}), \quad (34)$$

$$\frac{d[\text{CO}_2]}{dt} = u_{\text{CO}_2}, \quad (35)$$

$$\frac{d[\text{H}_2\text{O}]}{dt} = u_{\text{H}_2\text{O}}. \quad (36)$$

3.4.1 Heat of total combustion

Finally, we describe the temperature change due to the combustion reaction. The time change of the heat Q generated by the combustion reaction is

$$\begin{aligned} \frac{dQ}{dt} = & \Delta_f G_{\text{fuel}} u_{\text{fuel}} \\ & - \Delta_f G_{\text{C}} (au_{\text{fuel}} - u_{\text{CO}}) \\ & - \Delta_f G_{\text{H}} (bu_{\text{fuel}} - u_{\text{H}_2\text{O}}) \\ & - \Delta_f G_{\text{CO}} (u_{\text{CO}} - u_{\text{CO}_2}) \\ & - \Delta_f G_{\text{CO}_2} u_{\text{CO}_2} \\ & - \Delta_f G_{\text{H}_2\text{O}} u_{\text{H}_2\text{O}}, \end{aligned} \quad (37)$$

where $\Delta_f G$ is the standard Gibbs free energy of formation of the chemical species, which denotes the required energy

for the formation of chemical species from the standard state. Let c_p [J/(mol · K)] be the constant pressure specific heat of the particles. We use Eq.(38) to estimate $c_{p,\ell}$, and Eq.(39) to calculate c_p .

$$c_{p,\ell} = a + b\theta + c\theta^2 + d\theta^3 + e\theta^4 + \frac{f}{\theta^2}, \quad (38)$$

$$c_p = \frac{\sum_{\ell < \Theta} c_{p,\ell} d_{i,\ell}}{\sum_{\ell < \Theta} d_{i,\ell}}. \quad (39)$$

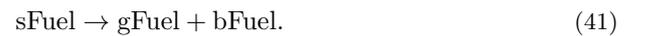
a to f is the constant coefficient. θ is $T/1000$. The temperature change in the particles dT/dt is determined as

$$\frac{dT}{dt} = \frac{1}{c_p \sum_{\ell < \Theta} d_{i,\ell}} \frac{dQ}{dt}. \quad (40)$$

3.5 Solid fuel heat decomposition

This section describes the extension of our combustion reaction model to solid fuel combustion. When solid fuel is burned, its decomposition generates fuel gas, which in turn reacts with oxygen and the flame appears. We model this phenomenon and consider the combustion of solid fuel. Note that to treat solid fuel consumption, solid deformation simulation need to be performed. However, this study focuses on flame simulation and does not treat solid fuel consumption.

Let sFuel be the solid fuel, gFuel be the fuel gas that is generated, and bFuel be the other products that are generated. The chemical reaction for thermal decomposition is assumed to be



The heat per unit mole generated by the Eq.(41) is $\Delta \mathbf{H}_{\text{dc}}$ [J/mol]. The reaction speed u_{dc} of Eq.(41) is determined as:

$$u_{\text{dc}} = \kappa_{\text{dc}} [\text{sFuel}]. \quad (42)$$

In this study, [sFuel] is a constant. Let ζ^g be the amount of fuel gas generated, T^{chS} be the temperature of the solid particle.

$$\frac{d\zeta^g}{dt} = u_{\text{dc}}, \quad (43)$$

$$\frac{dT^{\text{chS}}}{dt} = \frac{\Delta Q_{\text{dc}}}{C_p} = \frac{\Delta \mathbf{H}_{\text{dc}} u_{\text{dc}}}{d_{\text{sFuel}} c_p}, \quad (44)$$

where ΔQ_{dc} is the heat, C_p [J/K] is the heat capacity and d_{sFuel} [mol] is molar amount of sFuel. gFuel is assumed to be evenly distributed around the solid particles. The amount of fuel gas $d\zeta^g$ distributed to one gas particle is

$d\bar{\zeta}^g = d\zeta^g/n^{\text{amb}}$. n^{amb} is the number of gas particles in the vicinity of a solid particle. The combustion reaction model described in the previous section was applied for the combustion of fuel gas distributed in the gas particles.

3.6 Rendering

3.6.1 Flame color

In this section, we describe the determination of the color of the particles from the values obtained by the simulation. The color of the flame is determined using black body radiation of the soot and the color emitted by the reaction intermediate generated when the fuel molecules are decomposed.

First, we describe the color of the black body radiation of the soot. The surface area of the soot is assumed proportional to $[C]$. Let the surface area of a coffin be $S_{\text{soot}} = \varphi[C]$, where φ is the proportional constant. The spectral radiant flux $I_T(\lambda, T)$ due to the black body radiation of soot is given by the Planck's law:

$$I_T(\lambda, T) = \alpha S_{\text{soot}} \frac{8\pi hc}{\lambda^5} \frac{1}{\exp\left(\frac{hc}{\lambda K_b T} - 1\right)}, \quad (45)$$

$$= \alpha S_{\text{soot}} \dot{I}_T, \quad (46)$$

where λ , c , K_b , and h are the wavelength, speed of light, Boltzmann's constant, and Planck's constant, respectively. α is the user-defined parameter that determines the blackbody radiation. Equation (46) provides the intensity of each wavelength at a given temperature. Using the color matching function $\bar{\mathbf{w}}(\lambda) = (\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda))$, the relationship between the color values $\mathbf{W}_T = (X_T, Y_T, Z_T)$ in the XYZ color system and spectral radiant flux is

$$\mathbf{W}_T = \alpha S_{\text{soot}} \int \dot{I}_T(\lambda, T) \bar{\mathbf{w}}(\lambda) d\lambda. \quad (47)$$

Because a large amount of integration calculation is required, we create a table in advance and refer to the temperature in the table to obtain the values for the simulation.

Next, we describe the color due to the luminescence of the reaction intermediate. The luminescence of the reaction intermediate is composed mainly of the 430 nm emission spectrum of methylidene radical (CH), 510 nm emission spectrum of diatomic carbon (C_2), and 460 nm emission spectrum of carbon monoxide. The emission of methylidene radicals and diatomic carbon is confirmed when the fuel decomposes, and the emission of carbon monoxide is confirmed when it burns. Accordingly, we model the spectral flux I_r owing to the emission of the reaction intermediate as

$$I_r = \beta u_{\text{fuel}} N_A \frac{hc}{\lambda} (\delta(430 \text{ nm}) + \delta(510 \text{ nm})) + \gamma u_{\text{CO}_2} N_A \frac{hc}{\lambda} \delta(460 \text{ nm}). \quad (48)$$

where δ is the Dirac delta function and N_A is the Avogadro number. β and γ are user-defined parameters that determine the intensity of the light emission due to fuel decomposition and carbon monoxide combustion, respectively. As in the case of black-body radiation, color matching functions $\bar{\mathbf{w}}(\cdot)$ are used to obtain the color values \mathbf{W}_r in the XYZ color system as

$$\mathbf{W}_r = \beta u_{\text{fuel}} N_A \frac{hc}{\lambda_{430 \text{ nm}}} \bar{\mathbf{w}}(430 \text{ nm}) + \beta u_{\text{fuel}} N_A \frac{hc}{\lambda_{510 \text{ nm}}} \bar{\mathbf{w}}(510 \text{ nm}) + \gamma u_{\text{CO}_2} N_A \frac{hc}{\lambda_{460 \text{ nm}}} \bar{\mathbf{w}}(460 \text{ nm}). \quad (49)$$

The particle color value $\mathbf{W} = (X, Y, Z)$ is:

$$\mathbf{W} = \frac{\mathbf{W}_T + \mathbf{W}_r}{\|\mathbf{W}_T + \mathbf{W}_r\|_1 + H}, \quad (50)$$

where $\|\cdot\|_1$ denotes the L1 norm. Because \mathbf{W} is a normalized value, the synthesis of the two lights is given by $\mathbf{W}_T + \mathbf{W}_r / \|\mathbf{W}_T + \mathbf{W}_r\|_1$. In addition, the effect of ambient light on fire color is expressed by adding H to the denominator. $H = \varrho \varepsilon$ is the intensity of the ambient light. ϱ is the proportional constant, and ε is the user-defined parameter that determines the intensity of the ambient light. Finally, we convert the particle color \mathbf{W} into the sRGB color system $\mathbf{U} = (R, G, B)$. The user-defined parameter was made changeable during the simulation as a light intensity $\mathbf{L} = (\alpha, \beta, \gamma, \varepsilon)$. The user can change the color of the flame while watching the simulation result by adjusting \mathbf{L} .

3.6.2 Visual particles

To improve the quality of the image obtained by the simulation, it is necessary to reduce the radius of the particle that is to be positioned and to perform a more detailed simulation. However, a decrease in the particle radius increases the number of particles required for simulation, increasing the calculation time and decreasing the interactivity. Therefore, instead of increasing the number of particles used in the simulation, the visual particles were placed in the area where combustion reactions are likely to occur in the simulation space. This is the method used by Macklin et al.¹⁵⁾ to render smoke in CG.

To locate the drawn particles only in the area where the combustion reaction occurs, the visual particles are positioned in the area where fuel is supplied. The velocity

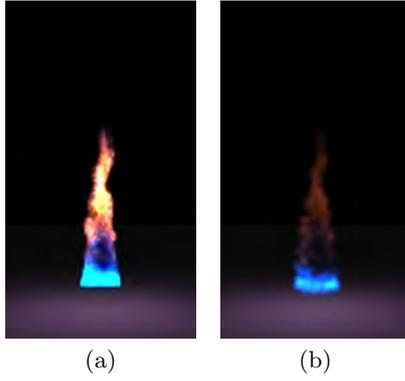


Fig. 3 Rendering results; (a) Visual particles, (b) Computational particles

Table 1 Particle i 's physical quantity

| Quantity | Notation |
|--------------------------|--------------------|
| mass | m_i |
| density | ρ_i |
| rest density | ρ_i^0 |
| amount of species ℓ | $d_{i,\ell}$ |
| temperature | T_i |
| ambient temperature | T_i^{amb} |
| PBD variable | λ_i |
| reaction rate of Eq.(22) | u_{CO_2} |
| reaction rate of Eq.(18) | u_{fuel} |

\mathbf{v}_i of the visual particle Ψ_i can be obtained from that of the computational particles around the visual particle using the equation

$$\mathbf{v}_i = \frac{\sum_{j < n} \mathbf{v}_j W(\mathbf{x}_i - \mathbf{x}_j, \xi)}{\sum_{j < n} W(\mathbf{x}_i - \mathbf{x}_j, \xi)}, \quad (51)$$

Using the same method, the color \mathbf{c}_i of the visual particle Ψ_i is obtained as

$$\mathbf{c}_i = \frac{\sum_{j < n} \mathbf{c}_j W(\mathbf{x}_i - \mathbf{x}_j, \xi)}{\sum_{j < n} W(\mathbf{x}_i - \mathbf{x}_j, \xi)}, \quad (52)$$

Figure 3 compares the rendering result of visual particles with that of computational particles, where the number of particles is the same. By arranging visual particles between computational particles, it is possible to express a finer flame while maintaining interactivity. In Fig. 3, the radius of visual particle is 0.35 times the radius of the computational particles.

4. Implementation

For ensuring interactivity, the proposed model is implemented on a GPU using CUDA. The pseudo code is shown in **Algorithm 1**.

First, the particle position and velocity are updated by the PBF method. Subsequently, the diffusion and chemical reaction equations are solved, and the temperature

Algorithm 1 Flame simulation

```

1: for all particle  $i$  do
2:   update velocity  $\mathbf{v}_i$  and position  $\mathbf{x}_i$  by PBF method
3: end for
4: for all particle  $i$  do
5:   calculate viscosity effect
6: end for
7: for all particle  $i$  do
8:   solve diffusion equation using Eq.(15)
9:   update each chemical species concentration and temperature
10: end for
11: for all particle  $i$  do
12:   compute decomposition reaction speed using Eq.(19) speed  $u$ 
13:   update each chemical species concentration
14:   compute combustion reaction speed  $u$  using Eq.(23), Eq.(24), and Eq.(27)
15:   update each chemical species concentration
16:   update temperature
17:   calculate ambient temperature
18: end for
19: for all particle  $i$  do
20:   calculate color and update
21:   update mass and default density
22: end for
23: for all visual particle  $i$  do
24:   calculate velocity and color  $\mathbf{v}_i$  and  $\mathbf{c}_i$  using Eq.(51) and Eq.(52)
25:   update position  $\mathbf{x}_i$ .
26: end for
    
```

and amount of chemical species contained in the particles are updated. Next, the velocity of the particles is modified using the updated temperature. Furthermore, the color of the particle is determined from the obtained particle temperature, amount of chemical species, and reaction rate. Finally, the velocity and color of the particles are determined for drawing.

We use the Poly6 kernel for W_{ij} and the Spiky kernel for ∇W_{ij} , as in Müller et al.¹⁸⁾. The following boundary condition is applied: particles that collide with the wall reverse their velocity in the normal direction. At the same time, the temperature and amount of chemical species are returned to the initial state. In this simulation, we used eight chemical species: oxygen, nitrogen, carbon dioxide, carbon monoxide, carbon atoms, hydrogen atoms, water, and fuels. The parameters of particle i are given in **Table 1**. The parameters used in this simulation are given in **Tables 2, 3, 4** and **5**. In Table 4, we show the Arrhenius parameter A and E_a of each chemical reactions. Decomposition reaction (Eqs. (18) and (41)) is not real reaction, thus there is no referenceable data. We implement Arrhenius parameter of Eqs. (18) and (41) as a tunable parameter. In Table 5, all coefficients (a-f) in Eq.(38) refer to Ref.19). In Table 5, D_0 of the chemical species without reference data was obtained using the

Table 2 Parameters

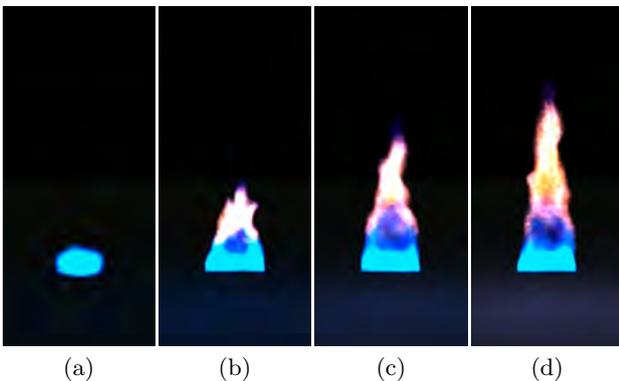
| Parameter | Notation | Value |
|-------------------------------|-------------|-----------------------|
| computational particle radius | r | 0.20 m |
| time step | Δt | 0.10 ms |
| gravity | g | 9.81 m/s ² |
| reference pressure | p_0 | 100000 Pa |
| initial temperature | T_0 | 20 °C |
| buoyancy constant | ϑ | 0.0051 |
| soot constant | φ | 1.0×10^{12} |
| ambient light constant | ϱ | 1.0×10^7 |
| heat absorbed | Q_{dc} | -100 kJ/mol |

Table 3 Standard Gibbs free energy of formation

| Species | Notation | Value [kJ/mol] |
|---------------------------------|--------------------------|----------------|
| CO ₂ | $\Delta_f G_{CO_2}$ | -394.39 |
| CO | $\Delta_f G_{CO}$ | -137.16 |
| C | $\Delta_f G_C$ | 671.20 |
| H | $\Delta_f G_H$ | 203.28 |
| H ₂ O | $\Delta_f G_{H_2O}$ | -228.14 |
| C ₂ H ₆ O | $\Delta_f G_{C_2H_6O}$ | -135.8 |
| C ₈ H ₁₈ | $\Delta_f G_{C_8H_{18}}$ | 6.8 |

Table 4 Arrhenius parameters

| Chemical reaction | A | E_a | Ref |
|-------------------|------------------------|---------------------|-----|
| Eq.(18) | 1.0×10^{10} | 130×10^3 | - |
| Eq.(21) | 3.745684×10^7 | 0.1×10^3 | 19) |
| Eq.(22) | 4.504×10^6 | 201×10^3 | 19) |
| Eq.(25) | 1.830688×10^8 | 67.55×10^3 | 19) |
| Eq.(41) | 1.0×10^5 | 150×10^3 | - |

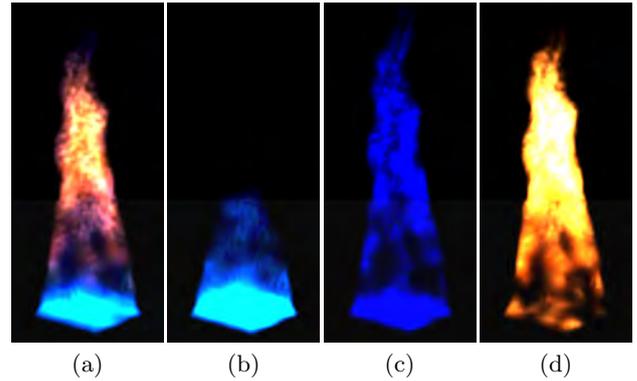
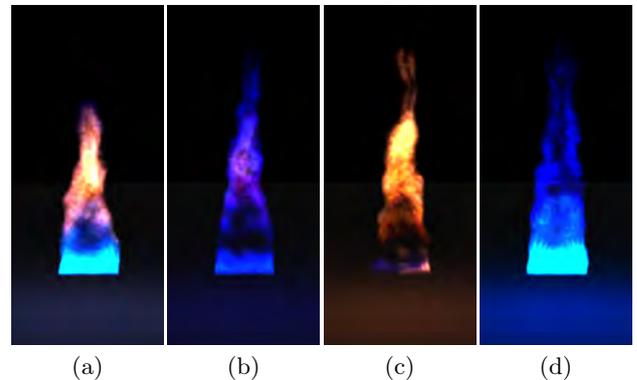

Fig. 4 Ignition experiment of ethanol by the proposed method; (a) A sample image immediately after ignition, and (b)–(d) Sample images after the passage of time

data of chemical species with similar atomic weights.

5. Results

In this section, we describe the results achieved by implementing and operating the proposed method. For performing the simulations, we used an NVIDIA GeForce GTX 1060 GPU, Intel Core i7 7700k CPU, and 16GB RAM. C++ was used as the development language, OpenGL 4.6 as the graphic library, and CUDA as the GPGPU environment.

Figure 4 shows the reproduction of the flame generation using the proposed method. We specified ethanol


Fig. 5 Split display of light source; (a) Original flame, (b) Highlighting the light derived from fuel decomposition, (c) Highlighting the light derived from the combustion of carbon monoxide, and (d) Highlighting the light from the black body radiation of the soot

Fig. 6 The color of the flame changed by manipulating the light intensity \mathbf{L} ; (a) $\mathbf{L} = (20, 0.4, 0.1, 0.1)$, (b) $\mathbf{L} = (1, 1, 0.02, 0.1)$, (c) $\mathbf{L} = (1, 1, 1, 1)$, and (d) $\mathbf{L} = (50, 1.25, 0, 0.1)$

(C₂H₆O) as the fuel and \mathbf{L} set to (20.0, 0.40, 0.10, 0.10). Figure 4(a) to Figure 4(d) shows the images after the fuel has been supplied and ignited.

Figure 5 shows the result when each light source is displayed separately. Figure 5(b) shows the turquoise blue light from the fuel decomposition and is found to occur mainly near the base of the flame. This result is consistent with the characteristics of the actual flame combustion phenomena. Figure 5(c) shows the pale blue light from the combustion of carbon monoxide that is particularly strong at the outer edge of the flame. Figure 5(d) shows the red and yellow light from the black body radiation of soot that is particularly strong at the tip of the flame. These results capture the characteristics of a real flame when the combustion is complete, which is mainly due to a large amount of oxygen at the outer edge of the flame; incomplete combustion occurs because of a lack of oxygen inside the flame, and a large amount of soot is generated.

As shown in Fig. 5, our simulation captures the ba-

Table 5 Coefficients of diffusion and specific heat

| Species | D_0 | Ref | a | b | c | d | e | f |
|---------------------------------|------------------------|-----|-----------|------------|-----------|-----------|---------|-----------|
| O ₂ | 0.21×10^{-4} | 20) | 30.03235 | 8.772972 | -3.988133 | 0.788313 | 0 | -0.741599 |
| N ₂ | 0.168×10^{-4} | 20) | 19.50583 | 19.88705 | -8.598535 | 1.369784 | 0 | 0.527601 |
| CO ₂ | 0.113×10^{-4} | 20) | 58.16639 | 2.720074 | -0.492289 | 0.038844 | 0 | -6.447293 |
| CO | 0.15×10^{-4} | - | 25.56759 | 6.096130 | 4.054656 | -2.671301 | 0 | 0.131021 |
| C | 0.181×10^{-4} | - | 21.17510 | -0.812428 | 0.448537 | -0.043256 | 0 | -0.013103 |
| H | 0.257×10^{-4} | - | 33.066178 | -11.363417 | 11.432816 | -2.772874 | 0 | -0.158558 |
| H ₂ O | 0.296×10^{-4} | 20) | 30.09200 | 6.832514 | 6.793435 | -2.534480 | 0 | 0.082139 |
| C ₂ H ₆ O | 0.12×10^{-4} | - | 4.9807 | 239.94 | -135.34 | 37.346 | -4.0291 | 0 |
| C ₈ H ₁₈ | 0.109×10^{-4} | - | -16.776 | 813.28 | -435.1 | 93.657 | 0 | 0 |
| CH ₄ | 0.10×10^{-4} | - | -0.703029 | 108.4774 | -42.52157 | 5.862788 | 0 | 0.678565 |

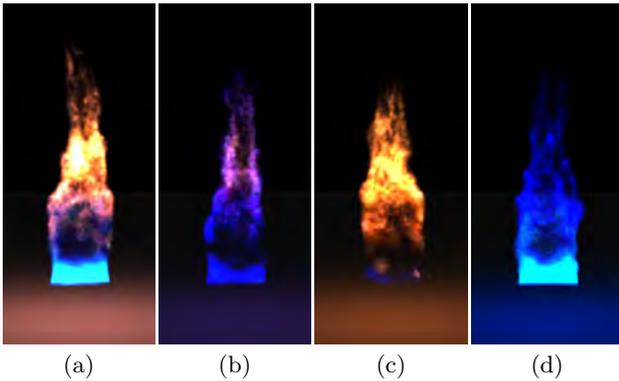


Fig. 7 Combustion simulation of octane; (a) $\mathbf{L} = (20, 0.4, 0.1, 0.1)$, (b) $\mathbf{L} = (1, 1, 0.02, 0.1)$, (c) $\mathbf{L} = (1, 1, 1, 1)$, and (d) $\mathbf{L} = (50, 1.25, 0, 0.1)$

sic characteristics of the flame combustion reaction and can easily create different types of flames by changing the light intensity \mathbf{L} (Fig. 6). In each column, the light intensity \mathbf{L} is equal. From the left column, \mathbf{L} is $(20, 0.4, 0.1, 0.1)$, $(1, 1, 0.02, 0.1)$, $(1, 1, 1, 1)$, and $(50, 1.25, 0, 0.1)$. Figure 6(a) shows the result for $\mathbf{L} = (20, 0.4, 0.1, 0.1)$. A flame with a blue base and red tip, as seen in candles, is generated. Figure 6(b) shows the result for $\mathbf{L} = (1, 1, 0.02, 0.1)$. A faint blue flame is produced, which is seen when alcohol is burned. Figure 6(c) is the result for $\mathbf{L} = (1, 1, 1, 1)$. A red flame is generated, which is common in bonfires. Figure 6(d) is the result for $\mathbf{L} = (50, 1.25, 0, 0.1)$. An intense blue-colored flame, seen mainly in gas burners, is produced.

Figure 7 shows the results when the fuel was changed from ethanol to octane (C₈H₁₈). The light intensity \mathbf{L} is the same as in Fig. 6. Although the flame is expanded sideways compared to the shape before the fuel was changed, there is no significant change in the color of the flame. The proposed combustion reaction model could not reproduce the changes in the flame color due to differences in the fuel.

To check whether our simulation can capture the features of real fire color, we compared the real images with

the simulation results. For comparison, we chose the bunsen burner fire. For the simulation, methane, which is the main fuel in the bunsen burner, was used. We lit the burner without supplying oxygen, and then increased the oxygen supply until the fire became completely blue. A video of the entire process was captured using a SONY FDR-AX700. $\mathbf{L} = (2, 1, 2, 0.1)$, and we add 1.2 m/s speed towards vertical upward in the fuel supply area. **Figure 8(a)** and Fig. 8(d) show that the images when only gas is supplied. Figure 8(b) and Fig. 8(e) capture the image when oxygen is supplied. Figure 8(c) and Fig. 8(f) are the cases where oxygen and gas are supplied and complete combustion occurs. In the real images, fire color changes with increasing supply of oxygen. Although the supply of oxygen varied, our simulation successfully captured this phenomenon by capturing the changes in the characteristics of real fire color, which is not possible using the conventional methods^{3),6),12)}.

In contrast, there is a difference between the real footage and simulation results in Fig. 8(c) and Fig. 8(f). In the real situation, oxygen was supplied at the center of the burner, but oxygen was supplied equally in the fuel supply area during the simulation. It is assumed that the difference in the distribution of oxygen supply caused a difference in the fire color, especially when the oxygen supply was high. By aligning the distribution of oxygen supply, more plausible simulation results can be achieved.

Figure 9 is an example of solid fuel combustion. Solid particles were placed near the tip of a match. The appearance of a fire from a lighter igniting the match is reproduced.

Figure 10 demonstrates changes in the flame caused by adding oxygen simultaneously with the fuel. Figure 10(a) shows the case where only fuel is used, Fig. 10(b) shows the case where a small amount of oxygen is supplied, and Fig. 10(c) shows the case where a large amount of oxygen is supplied. In all three figures,

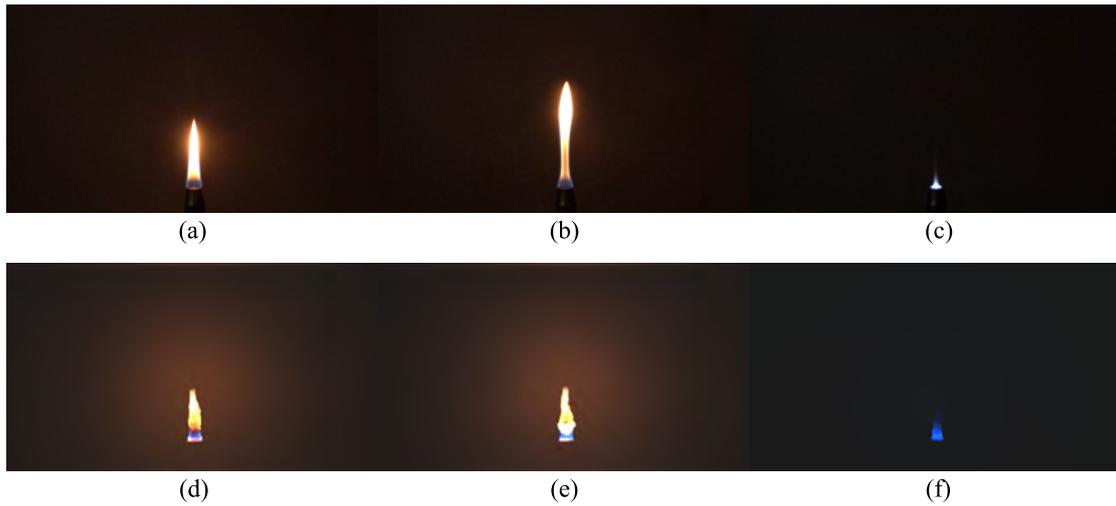


Fig. 8 Comparison of real iamges (upper) and corresponding simulation results(lower); (a)(d) Only gas is supplied, (b)(e) Oxgen is supplied, and (c)(f) Gas and oxygen are supplied

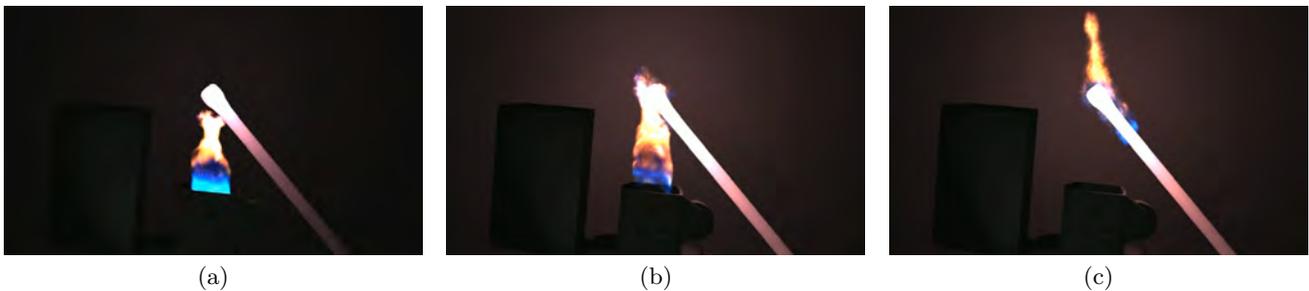


Fig. 9 Simulation of a burning match

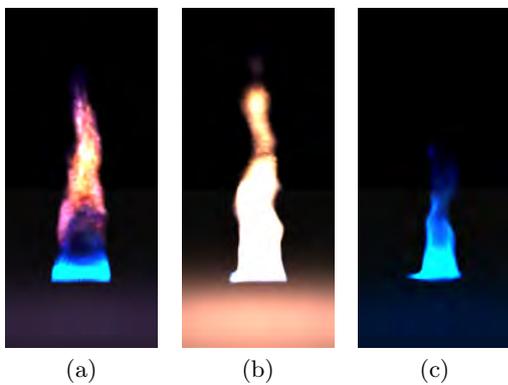


Fig. 10 Results of different supplying of oxygen to fuel; (a) When oxygen is not supplied, (b) When a small amount of oxygen is supplied, and (c) When a large amount of oxygen is supplied

$\mathbf{L} = (10, 0.5, 0.1, 0.1)$. The case where only fuel is used is similar to the real situation where lighter gas or a candle is burning. The supplied fuel diffuses and reacts with the oxygen originally present in the air. However, the case in which fuel and oxygen are added simultaneously is similar to the situation of combustion in a burner. By supplying a large amount of oxygen, the generation of soot can be eliminated completely as the fuel is burned completely. When a small amount of oxygen is added, the amount of

soot generated is clearly greater than when oxygen is not added. The addition of oxygen is considered to increase the likelihood that the reaction will proceed and cause a large amount of fuel decomposition, while the generation of water consumes most of the oxygen.

Figure 11 shows an example of a dragon's breath reproduced using the proposed method. The fuel is supplied and the initial velocity is added. Figure 11 (a)–(d) show the results of changing the light intensity \mathbf{L} for various flame colors.

The simulation time was used to evaluate the system. The relationship between the simulation time and the number of particles is summarized in **Table 6**. Our simulation operates at approximately 20 fps with 50,000 computational particles and 100,000 visual particles. According to Table 6, the computation time when using the proposed combustion model is reduced by approximately half to one-third than that by the constraint method.

6. Limitation

In our method, the combustion reaction is represented by the same reaction path as in Fms. (18), (21), (22), (25), (26) of the generalized fuel $C_aH_bO_c$. The values of

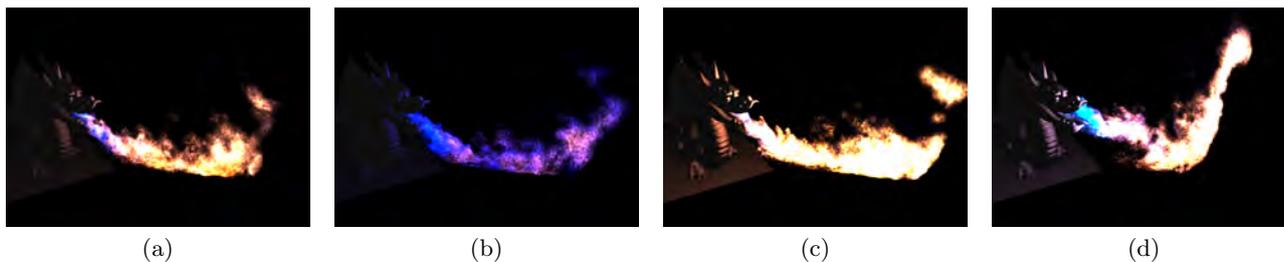

Fig. 11 Reproduction of a dragon's breath

Table 6 Relationship between computation time and number of particles

| Scene | Simulation space | Computational particles | Visual particles | Total time | Constraint Solver | Combustion and Diffusion |
|---------|------------------|-------------------------|------------------|------------|-------------------|--------------------------|
| Fig. 9 | 3m × 6m × 3m | 49,072 | 56,250 | 45.7 ms | 18.9 ms | 11.3 ms |
| Fig. 4 | 3m × 6m × 3m | 49,011 | 100,000 | 42.7 ms | 16.3 ms | 7.01 ms |
| Fig. 11 | 5m × 8m × 1m | 37,576 | 120,000 | 49.8 ms | 15.7 ms | 5.10 ms |

a, b, and c are changed by replacing the fuel, and the fire color becomes slightly different. In contrast, the combustion reaction path and the fire color should vary depending on the type of fuel. However, the proposed method does not cover the differences in the reaction paths. In addition, to handle large flames, a fluid simulation using many particles is required; however, interactivity is lost in this case. One solution is to position the particles along the vicinity of the flame, to simulate the flame in real time and shorten the calculation time.

7. Conclusion

In this paper, we proposed a simple combustion reaction model that can reproduce complex flame colors interactively. The combustion reaction, which involves thousands of chemical reactions, is modeled using five reactions. The blue and red colors of the flame can be simulated while maintaining interactivity. The calculation time of the simulation was measured, and it was confirmed that this simulation could operate at 20 fps or more if the number of calculated particles is approximately within a 50,000 limit. Thus, we can interactively change the color of the flame during the simulation. In contrast, by increasing the particle number, we can decrease the artifact. In addition, our proposed method adds the simulation of the combustion reaction and reactions to the fluid simulation of PBF, and therefore, it can be easily combined with physical simulation using other PBD models. For example, it is also possible to simulate the burning of a rigid body by combining the simulation with a rigid-body simulation produced by the PBD method. In the future, this study can be extended to simulate the combustion of a liquid by modeling its

evaporation.

Acknowledgement

This research was supported in part by JSPS KAKENHI under Grant 18H05010.

References

- 1) S. Sato, K. Mizutani, Y. Dobashi, T. Nishita, T. Yamamoto: "Feedback Control of Fire Simulation Based on Computational Fluid Dynamics", *Computer Animation and Virtual Worlds*, Vol. 28, No. 3-4, p. e1766 (2017).
- 2) J. Zhu, Y. Chang, E. Wu: "Realistic, Fast, and Controllable Simulation of Solid Combustion", *Computer Animation and Virtual Worlds*, Vol. 22, No. 2-3, pp. 125-132 (2011).
- 3) C. Horvath, W. Geiger: "Directable, High-Resolution Simulation of Fire on the GPU", *ACM Trans. on Graphics*, Vol. 28, No. 3, pp. 41:1-41:8 (2009).
- 4) Z. Melek: "Interactive Simulation of Fire, Burn and Decomposition", Ph.D. Thesis, Texas A&M University (2007).
- 5) A. Lamorlette, N. Foster: "Structural Modeling of Flames for a Production Environment", *ACM Trans. on Graphics*, Vol. 21, No. 3, pp. 729-735 (2002).
- 6) D. Q. Nguyen, R. Fedkiw, H. W. Jensen: "Physically Based Modeling and Animation of Fire", *ACM Trans. on Graphics*, Vol. 21, No. 3, pp. 721-728 (2002).
- 7) I. Ihm, B. Kang, D. Cha: "Animation of Reactive Gaseous Fluids through Chemical Kinetics", *Proc. of the 2004 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '04, pp. 203-212 (2004).
- 8) B. Kang, Y. Jang, I. Ihm: "Animation of Chemically Reactive Fluids Using a Hybrid Simulation Method", *Proc. of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '07, pp. 199-208 (2007).
- 9) M. Macklin, M. Müller: "Position Based Fluids", *ACM Trans. on Graphics*, Vol. 32, No. 4, pp. 104:1-104:12 (2013).
- 10) M. Faraday, *The Chemical History of a Candle*, Crowell (1957).
- 11) G. H. Markstein, *Nonsteady Flame Propagation*, Pergamon (1964).
- 12) J.-M. Hong, T. Shinar, R. Fedkiw: "Wrinkled Flames and Cellular Patterns", *ACM Trans. on Graphics*, Vol. 26, No. 3 (2007).
- 13) K. Min, D. Metaxas: "A Combustion-Based Technique for Fire

- Animation and Visualization”, The Visual Computer, Vol. 23, No. 9, pp. 679–687 (2007).
- 14) J. Stam, E. Fiume: “Depicting Fire and Other Gaseous Phenomena Using Diffusion Processes”, Proc. of ACM the 22nd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '95, pp. 129–136, (1995).
 - 15) M. Macklin, M. Müller, N. Chentanez, T. Kim: “Unified Particle Physics for Real-Time Applications”, ACM Trans. on Graphics, Vol. 33, No. 4, pp. 153:1–153:12 (2014).
 - 16) H. Schechter, R. Bridson: “Ghost SPH for Animating Water”, ACM Trans. on Graphics, Vol. 31, No. 4, pp. 61:1–61:8 (2012).
 - 17) N. M. Marinov: “A Detailed Chemical Kinetic Model for High Temperature Ethanol Oxidation”, International Journal of Chemical Kinetics, Vol. 31, No. 3, pp. 183–220 (1999).
 - 18) M. Müller, D. Charypar, M. Gross: “Particle-Based Fluid Simulation for Interactive Applications”, Proc. of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, SCA '03, pp. 154–159, Eurographics Association (2003).
 - 19) W. M. P.J. Linstrom, Nist Chemistry Webbook, Nist Standard Reference Database Number 69, <https://doi.org/10.18434/T4D303> (2020).
 - 20) Thermopedia, www.thermopedia.com/content/696/ (2020).

(Received July 2, 2020)
(Revised December 27, 2020)



Ryosuke UCHIKAWA

(Student Member)

He received the B.E. degrees from Osaka University, Toyonaka, Japan in 2019. His research interests include physically based simulation.



Hiroki ISHIZUKA

He received the Ph.D. degree in Integrated Design Engineering from Keio University in 2016. He is an Assistant Professor with the Graduate School of Engineering Science of Osaka University. His research interests include tactile displays and soft robotics.



Sei IKEDA

He received his Ph.D. degree from Nara Institute of Science and Technology, Ikoma, Japan in 2006. He is an Associate Professor with the Graduate School of Engineering Science of Osaka University. His research interests include human-computer interaction, mixed reality, and computer vision.



Yoshihiro KURODA

He received the Ph.D. degree in informatics from Kyoto University, Kyoto, Japan, in 2005. He has been a Professor with the Faculty of Engineering, Information and Systems at University of Tsukuba. His research interests include haptic interaction technologies and biomedical engineering.



Shunsuke YOSHIMOTO

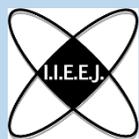
He received the Ph.D. degree in engineering from Osaka University, Toyonaka, Japan, in 2012. From 2020, he is a Lecturer with the Graduate School of Frontier Sciences, The University of Tokyo. His research interests include haptic engineering and biomedical instrumentation.



Osamu OSHIRO

He received the Ph.D. degree in engineering from Osaka University, Toyonaka, Japan, in 1990. He was a Research Engineer with the Sumitomo Metal Industries from 1990 to 1993 and was an Assistant/Associate Professor with the Nara Institute of Science and Technology from 1993 to 2003. Since 2003, he has been a Professor with the Graduate School of Engineering Science, Osaka University. His research interests include PBR, facing technology, modeling of distributed human system, etc.

3rd Call for Papers (Call for Late Breaking Papers)



The 7th IIEEJ International Conference on Image Electronics and Visual Computing 2021 (IEVC2021)

Shiretoko (Shari), Hokkaido, Japan / Sept. 8-11, 2021

<https://www.iieej.org/en/ievc2021/>

Purpose:

The International Conference on Image Electronics and Visual Computing 2021 (IEVC2021) will be held in Shiretoko (Shari), Hokkaido, Japan, on Sept. 8-11, 2021, as the 7th international academic event of the Institute of Image Electronics Engineers of Japan (IIEEJ) based on the great success of IEVC2007 held in Cairns, Australia, IEVC2010 held in Nice, France, IEVC2012 held in Kuching, Malaysia, IEVC2014 held in Koh Samui, Thailand, IEVC2017 in Danang, Vietnam, and IEVC2019 held in Bali, Indonesia. The conference aims to bring together researchers, engineers, developers, and students from various fields in both academia and industry for discussing the latest researches, standards, developments, implementations and application systems in all areas of image electronics and visual computing.



Conference venue:

The conference venue is Yumehall Shiretoko, Shari Town, Hokkaido, JAPAN.

Paper submission:

The official language is English, and authors should submit their papers as PDF through the online submission system. For details, visit the IEVC2021 official website.

(General papers submissions have already closed.)

Late Breaking Papers:

All suitably submitted papers for this category will be accepted for the conference. The authors must submit an abstract of which length is 1-2 pages, and select one from the following two types: 1) Technical papers or 2) Art/Demo papers. All the registered papers as late breaking papers will be published only in the USB proceedings of IEVC2021.

Important Dates

- | | |
|--|-----------------------------------|
| - Pre-Entry Submission (title, authors): | June 28, Monday, 2021 (extended) |
| - Abstract Submission (1-2 pages): | July 1, Thursday, 2021 (extended) |
| - Notification of Acceptance: | July 5, Monday, 2021 |
| - Camera-Ready Paper (1-2 pages): | July 9, Friday, 2021 (extended) |

Important information:

Online participation/presentation will be allowed for those who cannot come to the venue and make a presentation due to the COVID-19.



Call for Papers
Special Issue on
Image-related Technology for Realizing Immersive Media

IEEEJ Editorial Committee

Research on immersive media, which is an advanced system with such as all-sky video (360-degree video), VR, AR, and MR, is being actively conducted. In addition, the development of a head-mounted display (HMD) and the study of specific user interfaces are underway to fully demonstrate the appeal and the benefit of immersive media. In ISO / IEC, Coded Representation of Immersive Media (MPEG-I) has been studied for realizing immersive media in MPEG team, and though its standardization is being completed, further technologies are expected to be standardized. ITU-T SG16 is also promoting the recommendation of ultra-high presence live experience (ILE: Immersive Live Experience), and the momentum for systematization is increasing.

In this special issue, we are looking for a wide range of papers on elemental technologies that realize immersive media and system development papers on research to realize immersive high-presence systems by applying these technologies.

1. Topics covered include but not limited to

All-sky video (360-degree video), Free viewpoint Image, Point cloud, Light field, Holography, Head-mounted display (HMD), User interface, User experience, Usability, Interaction, Immersive content, VR, AR, MR, SR, xR, CG, Image processing, Image coding, Computer vision, Deep learning, MPEG-I, ILE

2. Treatment of papers

Submission paper style format and double-blind peer review process are the same as an ordinary contributed paper. If the number of accepted papers is less than the minimum number for the special issue, the acceptance paper will be published as an ordinary contributed paper. We ask for your understanding and cooperation.

3. Publication of Special Issue:

IEEEJ Transactions on Image Electronics and Visual Computing Vo.9, No.2 (December 2021)

4. Submission Deadline (**Extended**)

Monday, May 31, 2021 ⇒ Wednesday, June 30, 2021

5. Contact details for Inquires:

IEEEJ Office E-mail: hensyu@iieej.org

6. Online Submission URL: <http://www.editorialmanager.com/iieej/>

Call for Papers

Special Issue on CG & Image Processing Technologies Supporting and Expanding Human Creativities

IIEEJ Editorial Committee

The AI technology is expected to become a key technology for solving social problems and SDGs (Sustainable Development Goals), such as declining birthrate, aging population, shortage of labor, depopulation area, and so on, to which mature society, especially Japan is facing. The application scope of the technology in the field of image processing is expanding beyond the image generation and object recognition to include areas related to creativity, such as attribute conditioned image generation, super-resolution, image colorization, and line art coloring.

On the other hand, deep learning techniques don't always have good nature at interpretability and explainability of results required for creativity processing. Also, there is still issues to be addressed in AI technology, such as the ineffectiveness to the human-computer interface field by a simple application. So, technologies to support and expand the human creativity are gathering much concern, and quite promising in wide variety of applications.

This special issue of the paper targets various image-related technologies that support and expand creativities, and calls for papers and system development papers that cover not only applications of deep learning but also other technologies.

1. Topics covered include but not limited to
Image Processing, Image Recognition, Image Detection, Pattern Recognition,
Computer Graphics, Visualization, Binocular Vision, 3D image processing
Computer Vision, Big Data, Image Data Bases,
Machine Learning, Deep Learning, Understandability, Explainability,
Creativity, Usability, Interpretability, Human Interface and Interaction, User Experience, Ubiquitous,
Other related fundamental / application / systemized technologies.
2. Treatment of papers
Submission paper style format and double-blind peer review process are the same as an ordinary contributed paper. If the number of accepted papers is less than the minimum number for the special issue, the acceptance paper will be published as an ordinary contributed paper. We ask for your understanding and cooperation.
3. Publication of Special Issue:
IIEEJ Transactions on Image Electronics and Visual Computing Vo.10, No.1 (June 2022)
4. Submission Deadline:
Tuesday, November 30, 2021
5. Contact details for Inquires:
IIEEJ Office E-mail: hensyu@iieej.org
6. Online Submission URL: <http://www.editorialmanager.com/iieej/>

Guidance for Paper Submission

1. Submission of Papers

(1) Preparation before submission

- The authors should download “Guidance for Paper Submission” and “Style Format” from the “Academic Journals”, “English Journals” section of the Society website and prepare the paper for submission.
- Two versions of “Style Format” are available, TeX and MS Word. To reduce publishing costs and effort, use of TeX version is recommended.
- There are four categories of manuscripts as follows:
 - Ordinary paper: It should be a scholarly thesis on a unique study, development or investigation concerning image electronics engineering. This is an ordinary paper to propose new ideas and will be evaluated for novelty, utility, reliability and comprehensibility. As a general rule, the authors are requested to summarize a paper within eight pages.
 - Short paper: It is not yet a completed full paper, but instead a quick report of the partial result obtained at the preliminary stage as well as the knowledge obtained from the said result. As a general rule, the authors are requested to summarize a paper within four pages.
 - System development paper: It is a paper that is a combination of existing technology or it has its own novelty in addition to the novelty and utility of an ordinary paper, and the development results are superior to conventional methods or can be applied to other systems and demonstrates new knowledge. As a general rule, the authors are requested to summarize a paper within eight pages.
 - Data Paper: A summary of data obtained in the process of a survey, product development, test, application, and so on, which are the beneficial information for readers even though its novelty is not high. As a general rule, the authors are requested to summarize a paper within eight pages.
- To submit the manuscript for ordinary paper, short paper, system development paper, or data paper, at least one of the authors must be a member or a student member of the society.
- We prohibit the duplicate submission of a paper. If a full paper, short paper, system development paper, or data paper with the same content has been published or submitted to other open publishing forums by the same author, or at least one of the co-authors, it shall not be accepted as a rule. Open publishing forum implies internal or external books, magazines, bulletins and newsletters from government offices, schools, company organizations, etc. This regulation does not apply to a preliminary draft to be used at an annual meeting, seminar, symposium, conference, and lecture meeting of our society or other societies (including overseas societies). A paper that was once approved as a short paper and being submitted again as the full paper after completion is not regarded as a duplicate submission.

(2) Submission stage of a paper

- Delete all author information at the time of submission. However, deletion of reference information is the author’s discretion.
- At first, please register your name on the paper submission page of the following URL, and then log in again and fill in the necessary information. Use the “Style Format” to upload your manuscript. An applicant should use PDF format (converted from dvi of TeX or MS Word

format) for the manuscript. As a rule, charts (figures and tables) shall be inserted into the manuscript to use the “Style Format”. (a different type of data file, such as audio and video, can be uploaded at the same time for reference.)

<http://www.editorialmanager.com/iieej/>

- If you have any questions regarding the submission, please consult the editor at our office.

Contact:

Person in charge of editing

The Institute of Image Electronics Engineers of Japan

3-35-4-101, Arakawa, Arakawa-Ku, Tokyo 116-0002, Japan

E-mail: hensyu@iieej.org

Tel: +81-3-5615-2893, Fax: +81-3-5615-2894

2. Review of Papers and Procedures

(1) Review of a paper

- A manuscript is reviewed by professional reviewers of the relevant field. The reviewer will deem the paper “acceptance”, “conditionally acceptance” or “returned”. The applicant is notified of the result of the review by E-mail.

- Evaluation method

Ordinary papers are usually evaluated on the following criteria:

- ✓ Novelty: The contents of the paper are novel.
- ✓ Utility: The contents are useful for academic and industrial development.
- ✓ Reliability: The contents are considered trustworthy by the reviewer.
- ✓ Comprehensibility: The contents of the paper are clearly described and understood by the reviewer without misunderstanding.

Apart from the novelty and utility of an ordinary paper, a short paper can be evaluated by having a quickness on the research content and evaluated to have new knowledge with results even if that is partial or for specific use.

System development papers are evaluated based on the following criteria, apart from the novelty and utility of an ordinary paper.

- ✓ Novelty of system development: Even when integrated with existing technologies, the novelty of the combination, novelty of the system, novelty of knowledge obtained from the developed system, etc. are recognized as the novelty of the system.
- ✓ Utility of system development: It is comprehensively or partially superior compared to similar systems. Demonstrates a pioneering new application concept as a system. The combination has appropriate optimality for practical use. Demonstrates performance limitations and examples of performance of the system when put to practical use.

Apart from the novelty and utility of an ordinary paper, a data paper is considered novel if new deliverables of test, application and manufacturing, the introduction of new technology and proposals in the worksite have any priority, even though they are not necessarily original. Also, if the new deliverables are superior compared to the existing technology and are useful for academic and industrial development, they should be evaluated.

(2) Procedure after a review

- In case of acceptance, the author prepares a final manuscript (as mentioned in 3.).
- In the case of acceptance with comments by the reviewer, the author may revise the paper in consideration of the reviewer’s opinion and proceed to prepare the final manuscript (as

mentioned in 3.).

- In case of conditional acceptance, the author shall modify a paper based on the reviewer's requirements by a specified date (within 60 days), and submit the modified paper for approval. The corrected parts must be colored or underlined. A reply letter must be attached that carefully explains the corrections, assertions and future issues, etc., for all of the acceptance conditions.
- In case a paper is returned, the author cannot proceed to the next step. Please look at the reasons the reviewer lists for the return. We expect an applicant to try again after reviewing the content of the paper.

(3) Review request for a revised manuscript

- If you want to submit your paper after conditional acceptance, please submit the reply letter to the comments of the reviewers, and the revised manuscript with revision history to the submission site. Please note the designated date for submission. Revised manuscripts delayed more than the designated date be treated as new applications.
- In principle, a revised manuscript will be reviewed by the same reviewer. It is judged either acceptance or returned.
- After the judgment, please follow the same procedure as (2).

3. Submission of final manuscript for publication

(1) Submission of a final manuscript

- An author, who has received the notice of "Acceptance", will receive an email regarding the creation of the final manuscript. The author shall prepare a complete set of the final manuscript (electronic data) following the instructions given and send it to the office by the designated date.
- The final manuscript shall contain a source file (TeX edition or MS Word version) and a PDF file, eps files for all drawings (including bmp, jpg, png), an eps file for author's photograph (eps or jpg file of more than 300 dpi with length and breadth ratio 3:2, upper part of the body) for authors' introduction. Please submit these in a compressed format, such as a zip file.
- In the final manuscript, write the name of the authors, name of an organizations, introduction of authors, and if necessary, an appreciation acknowledgment. (cancel macros in the Style file)
- An author whose paper is accepted shall pay a page charge before publishing. It is the author's decision to purchase offprints. (ref. page charge and offprint price information)

(2) Galley print proof

- The author is requested to check the galley (hard copy) a couple of weeks before the paper is published in the journal. Please check the galley by the designated date (within one week). After making any corrections, scan the data and prepare a PDF file, and send it to our office by email. At that time, fill in the Offprint Purchase Slip and Copyright Form and return the scanned data to our office in PDF file form.
- In principle, the copyrights of all articles published in our journal, including electronic form, belong to our society.
- You can download the Offprint Purchase Slip and the Copyright Form from the journal on our homepage. (ref. Attachment 2: Offprint Purchase Slip, Attachment 3: Copyright Form)

(3) Publication

- After final proofreading, a paper is published in the Academic journal or English transaction (both in electronic format) and will also be posted on our homepage.

Editor in Chief: Mei Kodama
The Institute of Image Electronics Engineers of Japan
3-35-4-101, Arakawa, Arakawa-ku, Tokyo 116-0002, Japan

Print: ISSN 2188-1898
Online: ISSN 2188-1901
CD-ROM: ISSN 2188-191x
©2021 IEEEJ