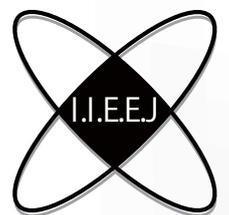


**IIEEJ Transactions on
Image Electronics and
Visual Computing**

Vol. 10, No. 2 2022



The Institute of Image Electronics Engineers of Japan

Editor in Chief

Osamu UCHIDA (Tokai University)

Vice Editors in Chief

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)

Mei KODAMA (Hiroshima University)

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 (FUJIFILM Business Innovation Corp.)

Yuki IGARASHI (Ochanomizu University)

Tomokazu ISHIKAWA (Toyo University)

Masahiro ISHIKAWA (Saitama Medical University)

Naoto KAWAMURA (Canon OB)

Shunichi KIMURA (FUJIFILM Business Innovation Corp.)

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.)

Tomooki MORIYA (Tokyo Denki University)

Paramesran RAVEENDRAN (University of Malaya)

Kaisei SAKURAI (DWANGO Co., Ltd.)

Koki SATO (Shonan Institute of Technology)

Syuhei SATO (Hosei University)

Masanori SEKINO (FUJIFILM Business Innovation Corp.)

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)

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.10 No.2 December 2022
CONTENTS**

Contributed Papers

- 141** Audio-Visual Judgement of In-Play or Out-of-Play in Volleyball Match Videos Reina ISHIKAWA, Ryo HACHIUMA, Hideo SAITO, Ryo FUJIWARA, Shoji YACHIDA
- 150** Multi-Sized Particle Sampling Method Based on Porosity Optimization in 2D Space Xu WANG, Makoto FUJISAWA, Masahiko MIKAWA
- 162** Real-Time Rendering of Oil Film with Flexible Properties Subroto Prasetyo HUDIONO, Tomoya ITO, Yuriko TAKESHIMA, Tsukasa KIKUCHI
- 170** Eyeliner Looked “Attractive” by Eyelids Shape Using Eyeliner Models Rena OKURI, Shuhei KODAMA, Tokiichiro TAKAHASHI

Guide for Authors

- 181** Guidance for Paper Submission

Audio-Visual Judgement of In-Play or Out-of-Play in Volleyball Match VideosReina ISHIKAWA[†], Ryo HACHIUMA[†], Hideo SAITO[†] (*Member*), Ryo FUJIWARA^{††}, Shoji YACHIDA^{††}[†] Keio University, Japan, ^{††} NEC, Japan

<Summary> Utilizing deep learning in sports match analysis is very promising for efficiency, and there have been many studies focusing on volleyball. However, previous studies have used only the frames where the ball is alive (*i.e.*, in-play frames), and recent data analysis methodologies rely on visual data, which are highly dependent on the position and angle of the camera relative to the court. As a result, these methods require a large dataset of images taken from various angles to improve accuracy. To efficiently extract and analyze the play as it is happening, we propose a model that distinguishes between in-play and out-of-play by combining visual data and audio data of volleyball match videos using late fusion. To investigate the effectiveness of the proposed model, we perform Grad-CAM visualization to determine which pixels the proposed model is focusing on.

Keywords: volleyball analysis, binary classification, audio-visual

1. Introduction

Sports analysis is beneficial to understand the strengths and conditions of a team and each player therein. By analyzing and quantifying the players' conditions in real-time during a match, the coach can give appropriate advice to the players or replace them with players who are better suited to the situation. Furthermore, by reviewing the analysis after a match, players can identify their weaknesses and use that information to develop effective training plans and new team tactics.

Conventionally, sports analysis had been implemented manually by specialists who are referred to as sports analysts. However, neural networks have appeared, and the idea of integrating sports analysis with neural networks has the potential power to remove conventional manual effort and limitations. This integration also enables a more efficient and detailed analysis. In particular, the automation of analysis in team sports such as volleyball and basketball, which are difficult to analyze manually, is in great demand given its ability to improve the efficiency and accuracy of the work. For example, individual or group action recognition¹⁾⁻⁴⁾ and player detection and tracking studies^{5),6)} provide an es-

sential foundation for analysis. Azar *et al.*¹⁾ proposed a method that undertakes both individual and group action recognition tasks to reflect individual-level actions in group-action recognition.

The problem with the conventional studies referenced above is that they use datasets that are supposed to contain in play data, so it is assumed that the out-of-play frames are removed beforehand. In this paper, by *in play*, we mean that *the ball is alive*, that is, it is in the time period when a referee blows the whistle to start the play to when the next whistle is blown. On the other hand, *out of play* means that *the ball is dead*, which refers to a period other than in-play. As volleyball is characterized by the fact that the out-of-play time is longer than the in-play time throughout the match, eliminating out-of-play time can lead to more efficient analysis. In this context, the automatic determination of whether the ball is in play is highly important for the downstream volleyball video analysis tasks, such as action recognition and ball/player tracking.

Therefore, we present a method that categorizes the scene as in play or out of play from the volleyball recordings. As our method eliminates redundant moments

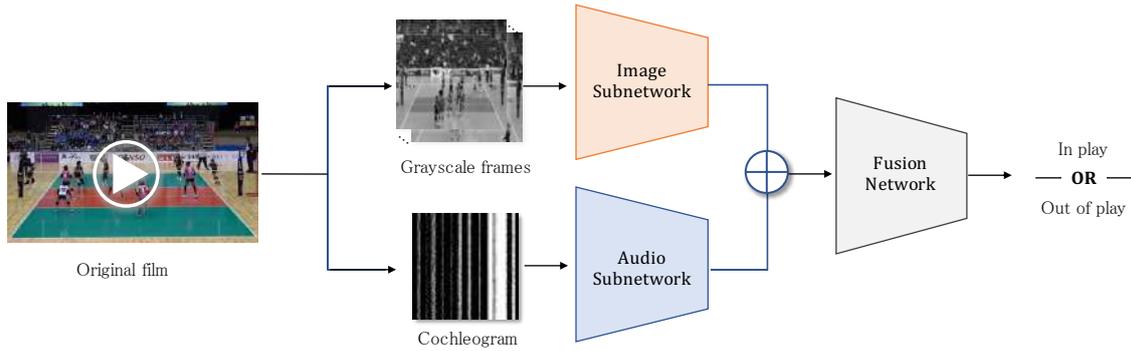


Fig. 1 Overview of the proposed method

(out-of-play frames) from the recorded video, it has the potential to speed up the volleyball analytics. By eliminating out-of-play scenes that are not useful in analytics, we expect that team staff, including players and coaches, can review their rallies in a shorter time. This elimination can let them smoothly analyze their states, such as their strengths or weakness. In addition, when we apply another process to the match video for further analysis, for example, using machine learning, eliminating redundant periods can reduce processing time.

We note that we are targeting volleyball staff, including volleyball analysts who are specialists in volleyball match analysis and coaches. Therefore in this paper, we assume to use raw videos taken by team members, and we do not use broadcast videos with replay scenes or overlaid graphics, such as scores and subtitles.

Moreover, most conventional prediction models¹⁾⁻⁴⁾ make decisions based only on visual data, and existing datasets contain only images in which the camera angle does not change significantly and the sight line is almost parallel to the centerline. However, such conventional methods cannot cope with significant changes in the viewpoint. Therefore, using only visual data requires a large amount of image data from various angles. In this study, we propose a method that uses audio data in addition to visual data. The audio data is considered to be less sensitive to changes in the angle of the viewpoint than the visual data and is thus expected to provide more robust judgment results. As far as we know, prior to this study, audio data has never been used in volleyball analysis.

Our contributions in this paper are as follows:

- We propose a framework that uses both audio and visual data to determine whether a given second in a volleyball game is in play or not using a convolutional neural network (CNN).
- We conduct experiments comparing the proposed method to the conventional method using either image or audio data. These experiments demonstrate how robust the proposed model is against changes in the seat occupancy rate or the camera position.
- We visualize the image subnetwork of the proposed model using Grad-CAM.

2. Related Work

Humans understand their surrounding environments in a multi-modal manner, using five senses: sight, touch, taste, smell, and hearing. When confidence in one modality is low, we can use the remaining modalities to make decisions; in other words, they can compensate for one other. These biological structures have recently inspired the investigation of multi-modal sensor usage in various fields. Here we introduce some previous works that have used multi-modalities and have focused on methods that incorporate audio features and visual features, which are used in this paper.

Understanding the ground conditions on which an autonomous driving robot will operate is crucial for ensuring smooth movement^{7),8)}. For example, Zörn *et al.*⁹⁾ proposed a self-supervised semantic segmentation method that takes advantage of independent audio and visual feature extraction.

The audio-visual feature is also utilized in the field of sports analysis. For instance, Xiong *et al.*¹⁰⁾ proposed a

highlight extraction method that associates visual object (e.g., baseball catcher and soccer goal post) detection and audio markers. According to the authors of⁽¹⁰⁾, each of the visual markers and audio markers can compensate for the other's false detection, which is a similar idea to the one we explore here. Khan⁽¹¹⁾ proposed an audio-visual highlight generalization method focusing on broadcast sports videos. They take advantage of audio features, such as the cheering of the audience, to identify *exciting points* within the video. Similarly, Grade *et al.*⁽¹²⁾ also proposed a sports-type classification method that combines visual motion features and audio feature processed with mel frequency cepstral coefficients (MFCCs) and principal component analysis (PCA). They successfully classify three sports types: basketball, soccer, and volleyball. The audio-visual usage is similar to ours, but the difference is the target of the classification. The previous study classifies the sports type and does not further split a single video. On the other hand, ours classify one video into many parts.

Itazuri *et al.*⁽¹³⁾ proposed a method to extract rally scenes by scoring camera transition and sound information. They assume to take the court in a specific direction because they are thinking of using the videos which are open to the public, and the users of the game are the general public. In contrast, our method targets sports analysts in a team who use videos taken by team members from their bench or audience seats. One of our contributions is robustness to the change of the camera position due to the usage of both visual information and audio information. Therefore, our study is uniquely positioned in terms of utilizing audio-visual data to recognize whether segments are in play or not in play in volleyball matches.

3. Methodology

In this study, segments are defined as one second, and the goal is to categorize each segment as either *in play* or *out of play* (see **Fig. 1**). For visual data, 10 frames are sampled uniformly from the target segment, and each segment is resized and converted into $144 \times 144 \times 10$ grayscale images. Then, the visual input is created by

randomly cropping a 128×128 image from the original grayscale frame and combining the 10 grayscale images. To process a one-dimensional audio signal into a two-dimensional input signal, we adopted cochleogram. Though there are other audio pre-processing methods, such as Mel-frequency cepstral coefficients (MFCCs), we suppose pitch information kept with cochleogram is essential to differentiate the sound of hitting the ball and the sounds of cheers, for example, applause. For this purpose, each audio segment is converted into two-dimensional data of 64×64 using cochleogram and used as input.

The image and audio subnetworks are constructed based on the VGG16⁽¹⁵⁾. However, the number of layers are changed from the VGG16 according to the input size, and the fully connected layers are removed. Each kernel size is 3×3 , and a max-pooling layer follows a set of convolutional layers. The image subnetwork consists of six layers, and the audio subnetwork consists of five layers. In the end, both subnetworks output a feature on a $256 \times 2 \times 2$ scale. Each feature output from the image and audio subnetworks is reshaped into a 1024-dimensional vector, and those are added and then input into the fusion network, which consists of 3 fully connected layers with a 0.5 dropout rate and a ReLU activation function. The output of the fusion network is a two-dimensional vector. Each element of the vector indicates the score for each class (*i.e.*, in play or out of play), and the final judgment is made based on which class has the higher score value.

4. Experiment

4.1 Dataset

In this study, we created a dataset from eight variety volleyball match videos to evaluate the proposed method. A sample image from each video is shown in **Fig. 2**. All of the volleyball match videos used in this paper were captured by a staff of the volleyball team playing from a static position. Among the videos taken from angles in #1–#8 of Fig. 2, a video from a particularly different angle is displayed in #4. The seat occupancy rates were 100% in the 1st, 2nd, 3rd, and 4th videos, 50% in the 5th and 6th videos, and almost 0%



Fig. 2 Sample images from each video

except for people affiliated with the team in the 7th and 8th videos. We manually annotate the approximated occupancy rates for videos.

Each video is segmented into 1-second intervals, and the image and audio for each segment are preprocessed as explained in Sec. 3. before they are put into each sub-network. **Table 1** shows both the length and the number of Segments in each video. To create the ground-truth labels, we manually labeled only the 4th video in Fig. 1. As for the remaining videos, we utilized the labels created by volleyball analysts in real-time while the match was going on. The part we want to extract as in-play is the period from when the 1st referee blows a whistle of starting a rally to when the referee blows the next whistle. However, in volleyball, there is a rule that a server has to carry out a service hit within 8 seconds after the 1st referee’s whistle. In other words, there may be an up to 8 seconds gap between the starting whistle and the time when the server hits the ball. Because analysts record the approximate time when a player tosses or releases the ball and hits it, their recorded time may differ from the actual time of the whistle. To fill in the gaps as much as possible, we manually labeled the three videos based on the referees’ whistles, which we assumed was the correct labeling, and we decided to take a margin to ensure the highest level of the accuracy of the labeling based on the analysts’ data. Adding margins of 2 seconds for the start of in play and +5 seconds for the end resulted in a 95.09% match between manual labeling and analysts’ labeling. The total number of in-play seg-

Table 1 The seat occupancy rate and the length of and number of segments in each video

Video	Seat occupancy rate (%)	Length	# Segments
1	100	01:56:31	6991
2	100	01:41:23	6083
3	100	02:30:52	9052
4	100	00:26:39	1599
5	50	02:25:53	8753
6	50	01:48:37	6517
7	0	02:18:49	8329
8	0	02:07:13	7633

ments is 19,996, and the number of out-of-play segments is 34,961.

4.2 Dataset splitting

We divided the dataset we generated into separate training, validation, and test datasets with two versions. We prepared a *heavy* version and a *light* version for each splitting method according to the seat occupancy rate shown in Table 1. Because the *heavy* version was utilized to examine the impact of the seat occupancy rate, it includes all eight recordings. In contrast, since the *light* version was utilized to examine the impact of differences in the court orientation, it includes the recordings that have a 100% seat occupancy rate.

Random split We randomly assigned 80% of the dataset to be part of a training dataset, 10% to be part of a validation dataset, and 10% to be part of a test dataset.

Time series split A split for training is made such that 80% of the time series video is allocated from the

beginning. The test dataset is from the next 10%, and the validation dataset is from the remaining 10%. The number of data in each dataset is almost the same as in the random split.

Sequence split We divided the dataset by video unit as shown in Table 1. First, we separated all the datasets into three groups according to the seat occupancy rate shown in the same table: 100%, 50%, and 0%. Then, each of the heavy and light versions was divided in the following ways:

- To investigate the impact of the seat occupancy rate, we implement leave-one-out classifications, that is, one of the groups is assigned to the test dataset, and the remaining groups are assigned to the training and validation datasets.
- The 1st, 2nd, and 3rd videos, which have similar court orientations, were assigned to the training and validation datasets. The data from the 4th video, which has the most variation in court orientation, was assigned to the test dataset. This data splitting is used to evaluate the effectiveness of audio data when the camera position relative to the court in the test video was different from the training dataset (see Subsec. 4.5.2).

Random splitting is the easiest split method for the model to learn the data distribution in each video, and it works as a benchmark. When the accuracy is high using the random split method and drops when using other splitting methods, the model is overfitting to short-time features. In this case, the model cannot classify the input data, even if it has been trained using the video with the same court orientation, and the method is never practically applicable. When the accuracy is high in the time series split and drops in the sequence split, the method can be applied only to other match videos using a similar court orientation. Our final goal is to achieve high accuracy even for unknown match videos, that is, to improve the accuracy by sequence segmentation.

4.3 Training details

Because the number of data per class is unbalanced, we adopt weighted cross-entropy for the loss function whose weights are hyperparameters. We set the weight

of the in-play class to 3 and the weight of the out-of-play class to 1. We train our model using the Adam optimizer¹⁶⁾ with a learning rate of 1e-6. The iteration number in the training time is set to 50 epochs for the *light* dataset (*i.e.*, the dataset that uses the 1st to the 4th videos) and 60 epochs for *heavy* dataset (*i.e.*, the dataset that uses the 1st to the 8th videos).

We used the model with the highest F-value against the validation data for testing. All of our implementations were based on the Pytorch framework,¹⁴⁾ and all the experiments were performed on a computer with an Intel(R) Xeon(R) W-3235 CPU and a Titan RTX GPU, 64GB memory.

4.4 Evaluation metric

We evaluated our model by calculating the accuracy, precision, recall, and F-score for each of the splitting methods, with *positive* indicating that the ball was alive and *negative* meaning that the ball was dead. The precision was determined based on the percentage of segments with positive labels among the segments the inference network judged as positive. The recall was determined by the percentage of the segments that the inference network judged as positive among those with positive labels. The F-score is the harmonic mean of the precision and the recall rates, which are defined as follows:

$$F = \frac{2}{\frac{1}{recall} + \frac{1}{precision}}. \quad (1)$$

4.5 Evaluation

To evaluate the effectiveness of our method, we created the *visual model* that uses only visual data and the *audio model* that uses only audio data, and we compared our method with those two models, respectively. The *visual model* receives only visual input and is similar to the image subnetwork. It has fully connected layers for classifying into in play or out of play in addition to the image subnetwork structure. Similarly, the *audio model* receives only audio input and has fully connected layers for classification after the audio subnetwork structure.

4.5.1 Seat occupancy rate

As the seat occupancy rate increases, the cheers from the audience get louder, which the audience cheering can

Table 2 Comparison between the *audio model* and our model using datasets with different seat occupancy rate

Dataset Split	Model	Accuracy↑	Precision↑	Recall↑	F-score↑
Random	Audio	0.804	0.685	0.845	0.757
	Ours	0.822	0.695	0.903	0.785
Time Series	Audio	0.792	0.668	0.850	0.748
	Ours	0.748	0.741	0.870	0.800
Sequence (test with 100% seat occupancy rate videos)	Audio	0.862	0.758	0.910	0.827
	Ours	0.807	0.736	0.732	0.734
Sequence (test with 50% seat occupancy rate videos)	Audio	0.789	0.662	0.854	0.746
	Ours	0.856	0.794	0.811	0.803
Sequence (test with 0% seat occupancy rate videos)	Audio	0.663	0.524	0.901	0.663
	Ours	0.719	0.576	0.889	0.699

affect the result of the classification task. Therefore, to investigate robustness against changes in the seat occupancy rate, we divided the dataset into three groups according to the seat occupancy rates, 100%, 50%, and 0%, and implemented leave-one-out training among these three groups. This is referred to as a *heavy sequence split*. This data split method requires the model to infer over the data with unknown seat occupancy rates. For instance, when we train our model using the data of 100% and 50% seat occupancy rates, we test our model using the data with a 0% seat occupancy rate. We also compared with those acquired using the *heavy random split* dataset and the *heavy time series split* dataset.

The results are shown in Table 2. With random split or time-series data split, our model surpassed the *audio model* in terms of precision, recall, and F-score. With sequence split, when the seat occupancy rate was 100%, the *audio model* outperforms ours in terms of both accuracy and F-score. This seems to result from the audience cheering significantly impacting the in-play and out-of-play classifications, and under the full occupancy rate condition, using only audio data can yield high accuracy. However, as the seat occupancy rate decreases, the precision rate of the *audio model* rapidly decreases because the audience cheering gets low. On the whole, our model demonstrates its robustness to the decrease of the seat occupancy rate. This result reveals that using an image subnetwork is effective.

4.5.2 Court orientation

In this subsection, we explore the effect of court orientation on our model using the 100% seat occupancy dataset, that is, the *light* version dataset. Among the

four videos, the 4th one has the most variation in court orientation of all the videos. We can observe from the result of Table 3 that with the sequence split dataset, where there were changes to the court orientation of the test data as compared to the random split or time series split datasets, the F-score and the accuracy dropped when we used the visual model. On the other hand, the proposed method, which uses both audio data and image data, resulted in better F-scores and accuracy with the sequence split than the visual model. Therefore, we can say that the proposed method is more robust to changes in the camera view than the visual model, and this is owing to the usage of the audio subnetwork.

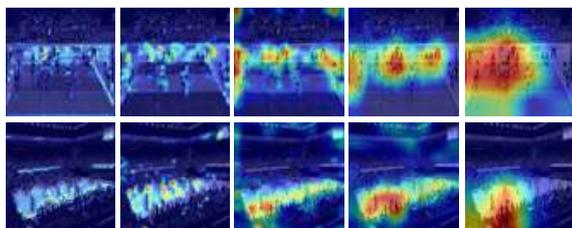
4.5.3 Visualization with GRAD-CAM

Subsection 4.5.2 demonstrated that using only visual data on its own may lead to erroneous decisions when the court orientation changes significantly, and the visual data may provide no information about the court orientation at all. With this in mind, in this subsection, we examine whether the visual network can capture information about the court and help the model make decisions. Specifically, we used Grad-CAM¹⁷⁾ to visualize the model trained on the sequence split dataset and examine which pixels are of interest in the visual input. Grad-CAM observes the gradients of inputs in each CNN layer to visualize which part of the input image is significant for predictions.

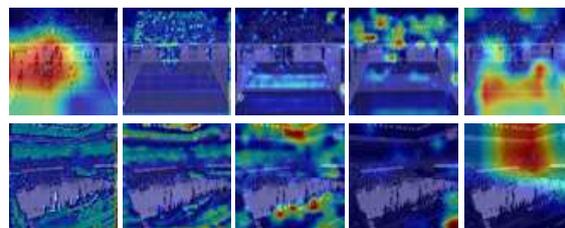
We show the visualization results in **Fig. 3**. From left to right, each image corresponds to the 1st to 5th layer in the proposed image subnetwork. Figure 3(a) represents the visualization of the layers by weighting the activation map based on the gradient of the in-play score for

Table 3 Evaluation using different court orientation video data

Dataset split	Model	Accuracy↑	Precision↑	Recall↑	F-score↑
Random	Visual	0.705	0.543	0.934	0.686
	Ours	0.796	0.640	0.937	0.761
Time Series	Visual	0.739	0.569	0.943	0.710
	Ours	0.811	0.657	0.927	0.769
Sequence	Visual	0.558	0.507	0.729	0.598
	Ours	0.736	0.704	0.716	0.710



(a) In-play visualization based on the in-play score



(b) Out-of-play visualization based on the out-of-play score

Fig. 3 Visualization samples created with Grad-CAM

an in-play image, and Fig. 3(b) shows the visualization of the layers by weighting the activation map based on the gradient of the out-of-play score for an out-of-play image. For each image group in (a) and (b), the upper result was generated using an image sample from the training data, and the lower one was generated using a sample from the test data. The figure shows that the inference network focuses on the players inside the court when drawing inferences from in-play images; In contrast, the network focuses on the unoccupied part of the court or the spectators when drawing inferences from out-of-play ones.

These results derive some considerations; To be detected as in-play, player positions can be informative; To be judged as out-of-play, the unoccupied part of the court or the move of the audience outside the court can be significant. Though the camera positions differ between the training dataset and the test dataset, we can see the network can detect the court position and the players inside the court. Because the featured part of the input image is the same regardless of the court orientation, the inference network can work, in other words, distinguish in-play or out-of-play without regard to the court orientation. Furthermore, since the attention areas are distinctly different for in-play versus out-of-play, image data is considered informative for in-play judg-

ment.

5. Discussion

Though our model demonstrated better results than *audio model* when the seat occupancy rate decreases, our model scored the lowest accuracy when using 0% seat occupancy rate videos. We assume this is because the cheering by the audience has a powerful impact on the in-play and the out-of-play classification. However, the most informative and invariable sounds are the referees' whistles and the sound of the plays inside the court for conducting the classification. Therefore, reducing the impact of the audience voices or extracting the whistles of the referees is expected to improve the robustness of the decrease in seat occupancy rate. Exploring these effects can be our future work.

In addition, the main limitation of this method is the labeling system. The labels for training are created by volleyball analysts or by team members during the volleyball match or created by the authors based on the whistle of the chief referee. The ground-truth boundaries between in-play and out-of-play created by analysts depend on the skill and judgment of the analysts. Moreover, as for the authors' labeling, when the sound of the whistle is blocked by cheers, picking out the whistle sound is difficult, so we need to label referencing the

move of the referees. For the above reasons, in both labeling methods, the boundaries between in-play scene and out-of-play scene can be blurred, and sometimes it is hard to define absolute boundaries. Therefore, we suppose the ground-truth labels can have a slight error, and it is difficult to achieve 100% accuracy.

6. Conclusion

In this paper, we presented an audio-visual binary classification method for determining whether a given second in a volleyball match video is in or out of play. Though the proposed method, which utilizes both audio and visual data, is quite simple, we showed it is robust to differences in the environment for each video since audio data and visual data can compensate for each other. Because we are targeting volleyball staff, including analysts and coaches, and we assume they use the practice or official match videos recorded by themselves, robustness against the changes in the viewing angle and changes in the seat occupancy rate is very beneficial.

We used Grad-CAM to verify which parts of the volleyball frame the model prioritizes. The Grad-CAM visualization revealed that the image subnetwork can recognize the inside and outside of the court, which is beneficial for making in-play or out-of-play decisions.

References

- 1) S. M. Azar, M. G. Atigh, A. Nickabadi, A. Alahi: "Convolutional Relational Machine for Group Activity Recognition," Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7884–7893 (2019).
- 2) J. Wu, L. Wang, L. Wang, J. Guo, G. Wu: "Learning Actor Relation Graphs for Group Activity Recognition," Proc. of the Computing Research Repository, pp. 9956–9966 (2019).
- 3) M. S. Ibrahim, S. Muralidharan, Z. Deng, A. Vahdat, G. Mori: "A Hierarchical Deep Temporal Model for Group Activity Recognition," Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1971–1980 (2016).
- 4) M. Qi, Y. Wang, J. Qin, A. Li, J. Luo, L. Van Gool: "stagnet: An Attentive Semantic RNN for Group Activity and Individual Action Recognition," Journal of IEEE Trans. on Circuits and Systems for Video Technology, Vol.30, No.2, pp. 549–565 (2020).
- 5) A. A. Sangüesa, G. Haro, C. Ballester: "Multi-Person Tracking by Multi-Scale Detection in Basketball Scenarios," Proc. of the Computing Research Repository (2019).
- 6) P. K. Santhosh, B. Kaarthick: "An Automated Player Detection and Tracking in Basketball Game," Journal of Computers, Materials & Continua, Vol.58, No.3, pp. 625–639 (2019).
- 7) A. Kurobe, Y. Nakajima, K. Kitani, H. Saito: "Audio-Visual Self-Supervised Terrain Type Recognition for Ground Mobile Platforms," Journal of IEEE Access, Vol. 9, pp. 29970–29979 (2021).
- 8) R. Ishikawa, R. Hachiuma, H. Saito: "Self-Supervised Audio-Visual Feature Learning for Single-Modal Incremental Terrain Type Clustering," Journal of IEEE Access, Vol. 9, pp. 64346–64357 (2021).
- 9) J. Zürn, W. Burgard, A. Valada: "Self-Supervised Visual Terrain Classification from Unsupervised Acoustic Feature Learning," Journal of IEEE Trans. on Robotics, Vol. 37, No. 2, pp. 466–481(2021).
- 10) Z. Xiong, R. Radhakrishnan, A. Divakaran, T. S. Huang: "Highlights extraction from sports video based on an audio-visual marker detection framework," Proc. of the IEEE International Conference on Multimedia and Expo (2005).
- 11) A. A. Khan, J. Shao, W. Ali, S. Tumrani: "Content-Aware Summarization of Broadcast Sports Videos: An Audio-Visual Feature Extraction Approach," Journal of Neural Process Letters, Vol.52, No.3, pp.1945–1968 (2020).
- 12) R. Grade, M. Abou-Zleikha, M. G. Christensen, T. B. Moeslund: "Audio-Visual Classification of Sports Types," Proc. of the IEEE International Conference on Computer Vision Workshop, pp. 768–773 (2015).
- 13) T. Itazuri, T. Fukusato, S. Yamaguchi, S. Morishima: "Court-Based Volleyball Video Summarization Focusing on Rally Scene," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 179–186 (2017).
- 14) A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejan, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala: "PyTorch: An Imperative Style, High-Performance Deep Learning Library," Proc of the Computing Research Repository (2019).
- 15) K. Simonyan, A. Zisserman: "Very Deep Convolutional Networks for Large-Scale Image Recognition," Proc. of the International Conference on Learning Representations (2015).
- 16) D. P. Kingma, J. Ba: "Adam: A Method for Stochastic Optimization," Proc. of the International Conference on Learning Representations (2015).
- 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 the IEEE International Conference on Computer Vision, pp. 618–626 (2017).

(Received August 19, 2022)

(Revised October 10, 2022)



Reina ISHIKAWA

She received her B.E. degree in information and computer science from Keio University, Japan, in 2021. She is currently pursuing her M.Sc.Eng. degree in science and technology at Keio University.



Ryo HACHIUMA

He received his B.E. and M.Sc.Eng. degree in information and computer science from Keio University, Japan, in 2016 and 2017, respectively. He received Ph.D. degree in science and technology with Keio University, Japan in 2021. He is currently working at Konicaminolta Japan, as a computer vision researcher.



Hideo SAITO (*Member*)

He received the Ph.D. degree in electrical engineering from Keio University, Japan, in 1992. Since 1992, he has been on the Faculty of Science and Technology, Keio University. From 1997 to 1999, he joined the Virtualized Reality Project as a Visiting Researcher at the Robotics Institute, Carnegie Mellon University. Since 2006, he has been a Full Professor at the Department of Information and Computer Science, Keio University.



Ryo FUJIWARA

He has been an analyst in the NEC Red Rockets Volleyball team since 2018, where he has contributed to the team training. He served as a support analyst for the Women's volleyball team of Japan in the Tokyo Olympics 2021. Since 2022, he has joined the Sports business promotional department of NEC Corporation, where he explores the sports DX area using new technologies and data analysis.



Shoji YACHIDA

He joined NEC Home Electronics, Ltd. in 1988. He joined NEC Corporation in 2000 and is now a senior researcher at the Visual Intelligence Research Laboratories of NEC. He has engaged the research for the high-speed vision sensing technology and its applications.

Multi-Sized Particle Sampling Method Based on Porosity Optimization in 2D Space

Xu WANG[†], Makoto FUJISAWA[†] (*Member*), Masahiko MIKAWA[†]

[†] University of Tsukuba

<Summary> This paper introduces a multi-sized particle sampling method within an arbitrary 2D shape using power tessellation. Our method aims to improve packing density as to sample as many particles as possible in a limited area. We propose a porosity-driven optimization technique to ensure no overlap between particles while increasing the packing density. With such properties, our method is applicable to physically-based simulations, such as the Discrete Element Method and its related framework. Additionally, this technology allows users to set the target particle size distribution by a predesigned cumulative distribution function and restrict the errors between 10% and 20% after the optimization. We demonstrate that our multi-sized particle sampling algorithms significantly improve packing density compared to Poisson disk sampling and SPH-based blue noise sampling.

Keywords: multi-sized particle sampling, porosity optimization, power tessellation, discrete element method

1. Introduction

Uniform particle sampling is a highly applicable technique in many graphics-related fields, such as image synthesis, rendering, geometry processing, and physically-based simulations. Notably, it has played an essential role in particle-based physics simulations in recent years. Meanwhile, several adaptive particle (multi-sized) sampling methods^{1)–3)} are also implemented into the fluid simulation. However, it is not well suited for simulating granular materials, such as sand, because slight overlap can destabilize the simulation. Therefore, we aim to achieve multi-sized particle sampling in arbitrary 2D polygons and have it available for using in stable physics-based sand simulations.

This paper proposes a multi-sized particle sampling algorithm using weighted centroidal power tessellation. It computes the particle distribution by minimizing the porosity in a pre-defined 2D polygon. Our contributions are as follows:

- We present a novel multi-sized particle sampling method based on minimum porosity. Our method ensures no overlapping occurs between particles and effectively exploits the limited space to improve sampling density.
- Our method supports users to give a target particle size distribution using a predesigned cumulative distribution function and limits the error between 10%

and 20% after the optimization.

- Multi-sized particles sampled by our method are available to the Discrete Element Method (DEM) and can realize stable physical simulations, whereas the particles sampled in the comparison experiments fail to perform stable physical simulations.

2. Related Work

2.1 Uniform particle sampling

Blue noise sampling is a well-known method used in computer graphics because of its ability to generate random uniform distributions. One of its patterns, Poisson disk sampling, has many rendering and geometry processing applications. Bridson⁴⁾ proposed a fast Poisson disk sampling method that is effective and more easily implemented in arbitrary dimensions. While this method has a wide application, it is difficult to control the sampling density and ensure no overlapping between particles. Traditionally, uniform particle sampling within a 2D polygon is mainly implemented by dividing the plane as the uniform grids (Voxelization in 3D⁵⁾) and generating a particle in each uniform grid. However, it will result in a particle distribution that does not fit the boundary shape. As a solution to this problem, Schechter et al.⁶⁾ proposed a relaxation-based sampling method to smoothen the particles on the boundary. Later, Jiang et al.⁷⁾ applied the Smoothed Particle Hydrodynamics (SPH) method to achieve blue noise sampling, which can also re-

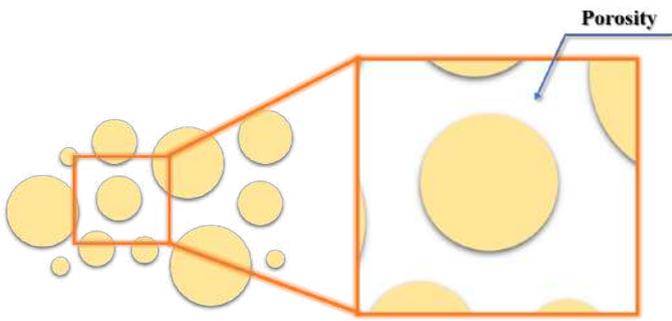


Fig. 1 Schematic diagram of porosity

alize uniform particle sampling by relaxing the boundary particles. Although these methods work well for uniform particle sampling, what we expect as multi-sized particle sampling and stable sand simulation cannot be fulfilled by simply extending their method.

2.2 Adaptive particle sampling

In order to achieve high-performance SPH-based fluid simulation, Adams et al.¹⁾ presented a dynamic sampling method, which can dynamically adjust the size of the sampled particles by computing the distance between the medial axis and fluid surface. Jiang et al.⁷⁾ integrated the adaptive particle sampling method and proposed an SPH-based blue noise sampling to treat multi-sized particles. Winchenbach et al.²⁾ proposed an adaptive incompressible SPH-based fluid simulation method and an improved framework³⁾ adapted to higher volume ratios for simulating low-viscosity turbulent flows and achieving the high-resolution rendering of fluid surfaces. However, since these algorithms are all SPH-based, there will be overlapping between the multi-sized particles, which make the multi-sized sand simulation unstable.

2.3 Voronoi diagram and centroidal voronoi tessellation

The Voronoi diagram has been widely used in the area of computer graphics. It has many applications in physics-based simulations of computer animation, such as power particles⁸⁾, adaptive staggered power particles⁹⁾, power diagram-based high-resolution adaptive liquids¹⁰⁾. Over time, several GPU-based algorithms diagram have been developed aiming to speed up the computation of the 3D Voronoi^{11)–14)}. Among these methods, centroidal Voronoi tessellation (CVT) for reaching a specific density distribution is also one of the critical topics related to the Voronoi diagram. CVT can be regarded as one of the relaxation-based sampling methods. It has two stages of computation: 1. Randomly generates the point set. 2. Using Lloyd iterations to optimize the loca-

tion of the points until convergence. However, directly using CVT will not cause the sampled points to have better properties of blue noise. Balzer et al.^{15),16)} proposed a capacity-constrained Voronoi tessellation, which can generate high-quality blue noise properties and perfectly fit a known density function. According to the feature that the energy function of CVT has second-order smoothness, Liu et al.¹⁷⁾ presented a faster convergence CVT method. Chen et al.¹⁸⁾ proposed a parallel algorithm for CVT on GPU and verified the algorithm's performance, which can obtain good blue noise characteristics and high-speed computing. All these methods described above have difficulty in extending to multi-sized particle sampling. On the other hand, CVT was used to study the hierarchical data visualization presented by Balzer et al.¹⁹⁾. They called this approach "Voronoi Treemaps". Later, Nocaj et al.²⁰⁾ introduced a faster, more straightforward, and resolution-independent method for computing Voronoi treemaps. While this approach can generate multi-sized particles, the user needs to pre-define a large amount of data since its computation is based on hierarchical data. Besides, sampled results are singular with the same data.

2.4 Physics based granular material simulation

Granular material simulation-related studies have a long history in the computer graphics community^{21)–25)}. Numerous works in recent years still focus on sand-water mixing simulation frameworks, such as SPH-based multi-phase multi-material flows^{26),27)} and material point method (MPM) based simulations of a mixture of sand and water^{28),29)}. Since the particle-particle approach (DEM) could better capture small-scale features at the particle level, several studies have also employed this method to make sandy-like animation, e.g., hybrid grains³⁰⁾ and seepage flow framework³¹⁾.

3. Porosity Optimization Based Approach

This section will introduce our porosity optimization-based method of how to sample multi-sized particles.

3.1 Controllable porosity

Generally, a rock or soil's porosity is measured by the fraction of the total soil volume occupied by pore space per unit volume. As shown in **Fig.1**, the porosity ϕ is defined as:

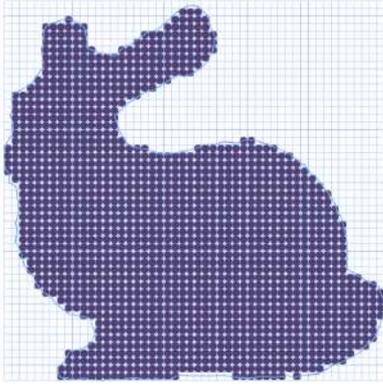


Fig. 2 Result of standard gridding approach based uniform particle sampling

$$\phi = 1 - \frac{V_{soil}}{V_{total}} \quad (1)$$

where V_{soil} is the volume of soil particles, and V_{total} presents a unit soil volume. The high ratio of soil particles in the unit volume will produce minor porosity. Hence we can improve the packing density within the unit volume by reducing the porosity and filling the limited space with more particles. Thanks to the controllability of the porosity, we can obtain higher packing densities and multi-sized particles by using a power diagram and porosity optimization.

3.2 Power diagram

There are n points $S = \{s_1, s_2, \dots, s_n\}$ (s_i called Voronoi site) in the plane, Voronoi diagram is made by dividing the plane into n regions $\mathcal{R} = \{\mathcal{R}(s_1), \mathcal{R}(s_2), \dots, \mathcal{R}(s_n)\}$ (each site $s_i \in S$ and its associated region $\mathcal{R}(s_i)$ are called Voronoi cell) and each region has only one point. Given an arbitrary 2D polygon as boundary Ω , Voronoi cell $\mathcal{R}(s_i)$ can be defined as:

$$\mathcal{R}(s_i) = \{p \in \Omega : \|p - s_i\| < \|p - s\| \text{ for each } s \in S - s_i\} \quad (2)$$

When assigning a weight w_i to each Voronoi site s_i , we call this Voronoi diagram as Power diagram which is defined as:

$$\mathcal{R}(s_i, w_i) = \{p \in \Omega : \|p - s_i\|^2 - w_i < \|p - s\|^2 - w \text{ for each } s \in S - s_i\} \quad (3)$$

When the weights w_i of all Voronoi site s_i are equal to zero, the power diagram is equivalent to a Voronoi diagram.

In contrast to the Voronoi diagram, the power diagram allows us to control the shape and size of each Voronoi

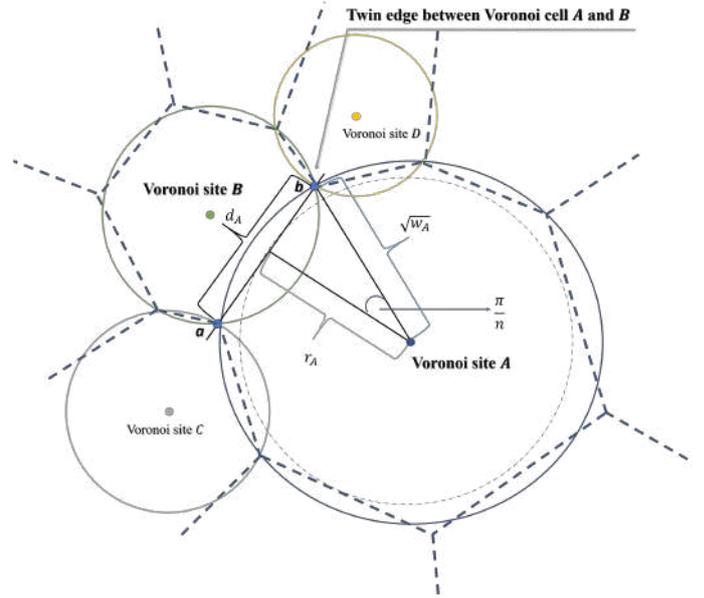


Fig. 3 A ideal case of the geometric relationship between the Voronoi cells and its weights

cell in a relatively simple way. It means that we can more readily manipulate the size of the sampled particles and the porosity of the entire space. For controlling the size of Voronoi cell, there are many studies related to the capacity-constrained method. In these studies, some methods combined power diagrams with a false position method to obtain the different sizes of Voronoi cells¹⁶⁾ or integrated with Newton's method³²⁾, L-BFGS method³³⁾. However, all these methods have a high computational cost and require optimization parameters tuning. As for our method, it is less complicated and more straightforward to control the size of the Voronoi cell. Though the precision is slightly inferior, it significantly improves packing density.

3.3 Requirements of multi-sized particle sampling

Using the standard gridding approach (**Fig.2**) as an example, when we generate particles in a unit cell, the minimum porosity is a fixed value of $1 - \frac{\pi}{4}$. It makes minimum porosity become a constant value and impossible to adjust dynamically.

According to the properties of the power diagram, the plane are divided into several Voronoi cells and each of them will be a convex polygon with a weight factor that can adjust its associated region. What inspires us is that it is possible to obtain sampling results with different porosity sizes by adjusting each Voronoi cell's weighting coefficients. During the Lloyd iteration, as changes in the weight coefficients of each Voronoi cell impact the shape and size of its associated region, it makes both the size of

particles and the pore size of its associated region changeable in each Voronoi cell. This property of the power diagram makes it possible to achieve multi-sized particle sampling within a given boundary. In order to achieve multi-sized particle sampling with a higher packing density, i.e., finding a minimum porosity sampling case, we summarize the requirements for the power diagram where the minimum porosity is obtainable, which we consider an ideal case:

Requirement 1: If each side of the Voronoi cell has the same length, then it is a regular polygon.

Requirement 2: If n Voronoi sites encircle a Voronoi site, its associated Voronoi cell is a regular n sided polygon.

Requirement 3: In the case of satisfying $n \rightarrow \infty$ and each side d of Voronoi cell has the same length, the shape of Voronoi cell is closer to the circle, porosity $\phi \rightarrow 0$.

Requirement 4: Each Voronoi site must be at the centroid of its associated Voronoi cell.

Notably, the ideal case described above is only a scenario envisioned for each Voronoi cell. It is impossible that the Voronoi cell can fully satisfy all the given four requirements locally or globally, but we can design the loss functions based on the requirements of an ideal case. Specifically, as the requirements are formulated for each Voronoi cell, the objective is to locally make each Voronoi cell meet the requirements and find a globally optimal solution.

3.4 Optimization

3.4.1 Objective function

To satisfy the requirements described in Section 3.3, we sketched **Fig.3** to represent an ideal case of the geometric relationship between the Voronoi cells and their weights. In this ideal case, Voronoi site A with B, C and D circling it satisfies the conditions in **Fig.4(b)**, which means that computing the length of twin edge \overline{ab} between Voronoi sites A and B is sufficient to know the weight of any site. Notably, it doesn't mean that we ignore the weight of the other Voronoi site but it has been taken into account in the computing power diagram (the power diagram will be reconstructed before updating the weights of Voronoi sites, as detailed in Algorithm 1). As for computing power diagrams, we mainly rely on the method

proposed by Aurenhammer³⁴). They bring a general d -dimensional power diagram up to compute with a $d+1$ dimensional convex hull. Briefly, they transform a 2D weighted Voronoi site $s \in S$ into a sphere with coordinates (x_s, y_s) and radius $r = \sqrt{w_s}$ mapped into the 3D plane, and then intersect with other transformed Voronoi sites on a 3D plane and finally apply an inverse transformation project onto the 2D plane (details can be found in Aurenhammer³⁴) or Nocaj et al.²⁰). The circles with the solid line in Fig.3 are the projection results, and their radii are the root of the weights of the associated Voronoi sites (i.e., the radius of Voronoi site A equal to $\sqrt{w_A}$).

Assuming that there exist n points around Voronoi site S_A , Ideally, the side length d_A of $\mathcal{R}(S_A)$ can be computed as follows:

$$d_A = 2\sqrt{w_A - r_A^2} \tag{4}$$

where w_A is the weight of Voronoi site S_A and r_A is the radius of Voronoi cell $\mathcal{R}(S_A)$'s maximum inscribed circle.

From the geometric properties of the power diagram, Equation 4 is only established in the case of Fig.4(b). When the weight of the Voronoi site is less than r_A^2 , it will produce the case (a) or case (c) in Fig.4, which means that we will use a penalty function ψ (Section 3.4.2) instead of shape loss function $\Phi_{\mathcal{P}}$ for preventing two sites from being too close or too far apart.

To satisfy requirements 1, 2, and 3 in Section 3.3, we focus on information about the adjacent Voronoi sites rather than the shape of the target Voronoi cell. It is because each edge of the target Voronoi cell is generated using the weights of the neighboring Voronoi sites. We consider updating the weight of the target Voronoi site for each iteration while the weights of the neighboring Voronoi sites are fixed. Consequently, to keep each edge of the target Voronoi cell of the same length as possible, the update tendency of the target Voronoi site weight can be computed by incorporating the weights of all neighboring Voronoi sites. One can define the loss function that controls the shape part of the porosity optimization as follows:

$$\Phi_{\mathcal{P}}(w, r) = \sum_{i \in S} \sum_{j \in S-i} |w_i - w_j + r_j^2 - r_i^2| \tag{5}$$

Meanwhile, we set a target particle radius distribution (Equation 10) in the initial stage to control the size of the sampled particles. According to this distribution function, we use the inverse transform method³⁵) to ensure that the generated target particle radii are uniformly dis-

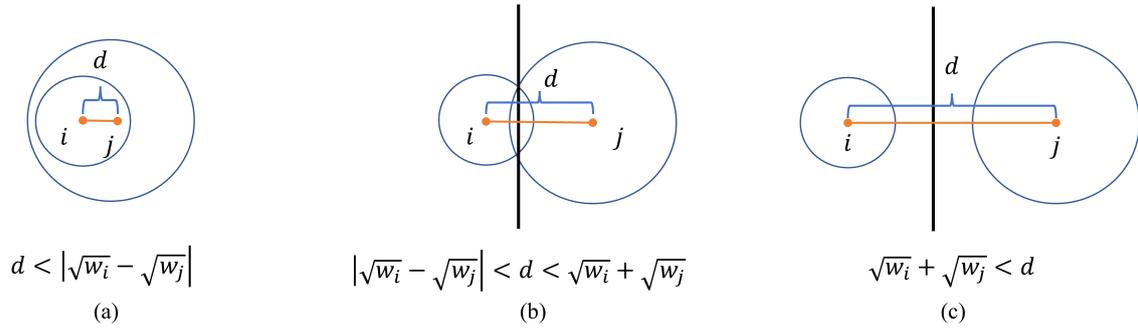


Fig. 4 Three possible cases between two neighboring Voronoi sites

tributed in each subinterval. Then we assign the generated target radius r_i to each Voronoi site s_i when generating the initial power diagram or dynamically appending particles (Section 3.4.3). Locally, the radius of the maximum inscribed circle in each Voronoi cell will approach the user's pre-set target radius. The global particle radius distribution obtained from multi-sized particle sampling will approximate the pre-defined particle radius distribution. Therefore, the area of each Voronoi cell in the ideal case (assume a regular n -sided polygon) can be computed by:

$$A_i^{\text{ideal}} = \frac{nr_i^2}{\tan \frac{\pi}{n}} \quad (6)$$

We re-generate the power diagram with new weights with each iterative computation, and each Voronoi site corresponds to a new Voronoi cell. In order to achieve the target area for each newly generated Voronoi cell, we define the loss function that controls the area part of the porosity optimization as follows:

$$\Phi_{\mathcal{A}}(A) = \sum_{i \in S} \frac{A_i^{\text{ideal}} - A_i^{\text{current}}}{A^{\text{total}}} \quad (7)$$

where A_i^{current} is the area of Voronoi site S_i associated region in the current iteration step, and A^{total} is the area of the boundary polygon Ω , which is the sum of all Voronoi cell areas. Noticeably, the A_i^{current} is not modified directly by the cost function, but it depends on the shape of the corresponding Voronoi cell in the updated power diagram. In addition, A_i^{current} is also used for computing the porosity when we append new particles.

With joint the shape loss term $\Phi_{\mathcal{P}}$ and area loss term $\Phi_{\mathcal{A}}$, the final loss function we minimise can be defined as:

$$\Phi(w, r, A) = \alpha \Phi_{\mathcal{P}}(w, r) + \beta \Phi_{\mathcal{A}}(A) \quad (8)$$

where α and β are two weighted coefficients for adjusting global optimization. We apply Lloyd iteration to the

power diagram to make each Voronoi cell meet the area requirement. Moreover, as the geometric relationship between the Voronoi cells and their weights is known, it allows us to optimize the shape of the Voronoi cells by adjusting the weights. Balzer and Deussen first proposed optimally meeting area requirements via Lloyd's method¹⁹⁾. It was followed by "Voronoi Treemaps",²⁰⁾ which also proved the practicability of this method. The core idea is that the weight of each Voronoi site is increased or decreased in proportion to the missing or redundant areas during Lloyd iteration. In our experiments, we found that this method is also applicable for controlling the shape of Voronoi cells, as shown in **Fig.5** for the optimization effect of porosity. It should note that we could perform L-BFGS method to optimize the power diagram^{32),33)}, but Lloyd's method is more straightforward, faster, and easier to implement. We would like to compare the impact of Lloyd's and L-BFGS methods on our experimental results in the future.

3.4.2 Constrained condition

While the Voronoi diagram ensures that each Voronoi cell only contains one Voronoi site, the power diagram may produce 3 cases, as shown in Fig.4, which shows that the Voronoi cell may contain zero or several Voronoi sites. Furthermore, the case of a Voronoi cell containing no Voronoi site or more than one Voronoi site will lead to the wrong results of our multi-sized particle sampling method. Our solution is not to perform the optimization described in Section 3.4.1 when the distance d_{ij} between Voronoi site i and j is greater than $\sqrt{w_i} + \sqrt{w_j}$ or less than $|\sqrt{w_i} - \sqrt{w_j}|$. Instead, we define a penalty function ψ for preventing two sites from being too close or too far apart, shown as follows:

$$\psi(w_i, d_{ij}, \sigma_k) = \sigma_k (d_{ij}^2 - w_i) \quad (9)$$

where $\sigma_k = k^2$ presents the penalty factor in k -th iter-

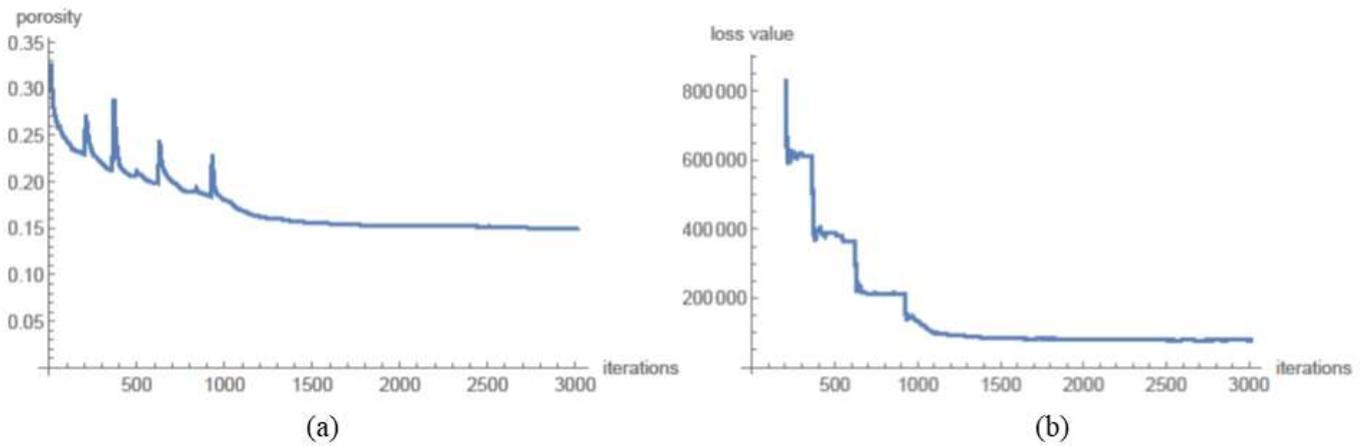


Fig. 5 The variation of porosity and loss value in the Bunny example

ation.

3.4.3 Dynamic particles appending

A fundamental issue with power diagram-based multi-sized particle sampling is that it requires searching for the optimal number of particles to fill the limited space. The reason is that generating a power diagram is needed to predetermine the number and weights of Voronoi sites. While the optimum number of particles for filling the entire space can be roughly estimated from the pre-defined particle distribution, the weights of each Voronoi site are difficult to compute in advance. To solve this problem, we propose a dynamic particle appending based on porosity. Specifically, we re-compute the porosity in the given boundary Ω at the end of each iteration and record it. Within the recorded data series, if the current iteration step is T_i , we will extract the porosity data change. Within a range from T_{i-k} to T_i (data length is k), we determine whether to append particles. We are using the least-squares method to fit the extracted porosity data and evaluate the slope of the fitted line. When it is lower than the threshold ϵ , we will append new particles according to the pre-defined particle radius distribution. We set $k = 20$ and $\epsilon = 1e-6$ in all our experiments and illustrate the process of dynamic particles appending method in Fig.6.

4. Implementation

This section will introduce the multi-sized particle sampling method described in Algorithm 1. The input to our algorithm is a boundary polygon Ω , composed of a series of 2D vertices and a list of target particle radii, and the list is generated from a user-defined particle radius dis-

tribution. The output is a power diagram consisting of sampled n Voronoi sites, where each site contains a particle radius. First, before the porosity optimization, we use the Lloyd iteration to ensure that each Voronoi site is located at the centroid of the Voronoi cell (line 22) followed by a search for neighboring Voronoi sites (line 4). We do not have to implement a specific neighbor search algorithm. The neighbor sites' information can be obtained from the QuickHull algorithm.

To address the 3 cases in the power diagram (Fig.4), the weights of the target and neighboring Voronoi sites determine whether to perform the objective function for porosity optimization or a penalty function (lines 6 to 12). Line 13 describes the degree to which the area of each Voronoi cell needs to be altered. Similarly, line 14 describes the degree of variation which is capable of making the shape of each Voronoi cell a regular n-sided polygon. Then we compute the values of the area and shape weights respectively (weights w_i^A and w_i^P are increased or decreased proportionally to $\Phi_{\mathcal{A}}^i$ and $\Phi_{\mathcal{P}}^i$) for each Voronoi cell that need to be updated for reaching the ideal case (lines 16 to 20). After updating the weighted factors, we need to recompute the power diagram and perform Lloyd iteration to make sure that each Voronoi site is located at the centroid of Voronoi cell (line 21 and line 22). Finally, we use "Dynamic Particles Appending Mechanism" to determine the convergence of the current porosity and whether the additional particles are needed (line 23). In all of our experiments, we set $\alpha = 1$ and $\beta = 1$ for a stable and effective optimization.

We have implemented our multi-sized particle sampling algorithm using C++ on AMD Ryzen 9 3900X CPU. In

Algorithm 1 Multi-Sized Particle Sampling Based on Porosity Optimization

Input: Boundary Polygon Ω , Target Particle Radius R_1, R_2, \dots, R_n
Output: Power Diagram $\mathcal{R}(S)$, Voronoi Sites $s_i \in S$ with a particle radius

```

1: repeat
2:    $\mathcal{R}(S) \leftarrow$  Compute Power Diagram
3:   for each  $s_i$  in  $S$  do
4:      $s_i$  Neighbor Searching( $n_i$ : Number of Neighbor Sites  $s_j \in S'_i$ )
5:      $\Phi_{\mathcal{P}}^i, \Phi_{\mathcal{P}}^{\text{total}} \leftarrow 0$ 
6:     for each  $s_j$  in  $S'_i$  do
7:       if  $(d_{ij} < |\sqrt{w_i} - \sqrt{w_j}|) \vee (\sqrt{w_i} + \sqrt{w_j} < d_{ij})$  then
8:          $\Phi_{\mathcal{P}}^i += \sigma_k(d_{ij}^2 - w_i)$ 
9:       else
10:         $\Phi_{\mathcal{P}}^i += \sum_j (\frac{\sum_j (w_j - r_j^2)}{n_i} - (w_j - r_j^2))$ 
11:      end if
12:    end for
13:     $\Phi_{\mathcal{P}}^{\text{total}} += \Phi_{\mathcal{P}}^i$ 
14:     $\Phi_{\mathcal{A}}^i = \frac{A_i^{\text{idea}} - A_i^{\text{current}}}{A_i^{\text{total}}}$  ( $A_i^{\text{idea}} = \frac{n_i r_i^2}{\tan \frac{\pi}{n_i}}$ )
15:     $\Phi_{\mathcal{P}}^i = \frac{\Phi_{\mathcal{P}}^i}{\Phi_{\mathcal{P}}^{\text{total}}}$ 
16:  end for
17:  for each  $s_i$  in  $S$  do
18:     $w_i^{\mathcal{A}} = 1 - (\Phi_{\mathcal{A}}^i - 1)^2$ 
19:     $w_i^{\mathcal{P}} = 1 - (\Phi_{\mathcal{P}}^i - 1)^2$ 
20:     $w_i += \alpha w_i^{\mathcal{A}} + \beta w_i^{\mathcal{P}}$ 
21:  end for
22:   $\mathcal{R}(S) \leftarrow$  Compute Power Diagram
23:  Performing Lloyd iteration
24:  Performing Dynamic Particles Appending Mechanism
25: until  $\Phi_{\text{global}} < \epsilon$ 
    
```

line 2 of Algorithm 1, we use the quick hull algorithm³⁶⁾ to compute the power diagram, which is stable, fast, and scalable for arbitrary dimensions. Additionally, we generate the straight skeleton³⁷⁾ in each Voronoi cell and compute the position and radius of the maximum inscribed circle in the Voronoi cell based on the straight skeleton as the output of Algorithm 1.

5. Results

In the experiment, we set up the particle radius distribution at initialization by using a cumulative distribution function as follows:

$$F_{\text{radius}}(r) = \begin{cases} 0.5 & \text{if } 20 < r \leq 30 \\ 0.4 & \text{if } 30 < r \leq 80 \\ 0.1 & \text{if } 80 < r \leq 150 \end{cases} \quad (10)$$

At initialization, we randomly generate 100 points within the boundary polygon Ω and set the maximum iterations to 3000.

Figure 7 shows the results of our multi-sized particle sampling method at the initial stage, 200th iterations, 400th iterations, 2000th iterations, and 3000th iterations. We also recorded the values of the global porosity

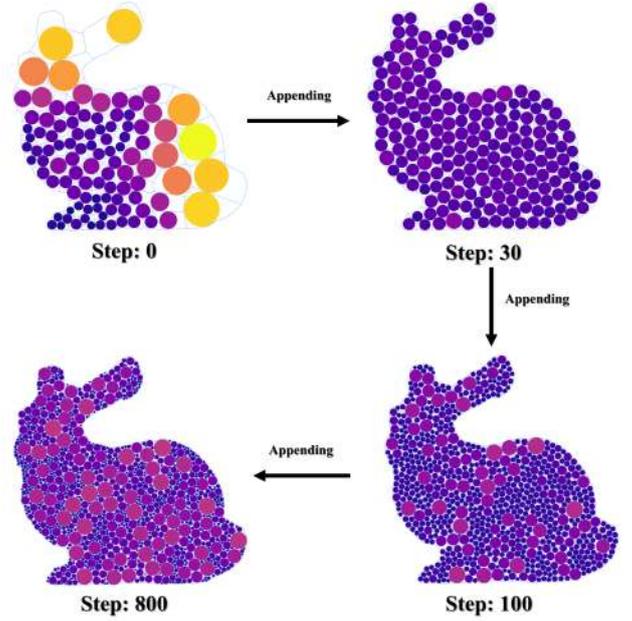


Fig. 6 Dynamic particles appending mechanism

(Fig.5(a)) and the loss value (Fig.5(b)) during 3000 iterations by taking Bunny as an example. It is worth noting here that in Fig.5(b), each time the loss value drops drastically, it is caused by triggering the dynamic particle appending mechanism.

Even though the porosity in Fig.5(a) also declines steadily when dynamic particle appending mechanism occurs, a sudden porosity increase occurs in the shortest time after each particle appending. This phenomenon is caused by the global porosity optimization having reached a steady state before adding the particles. Nevertheless, after appending new particles, since the shape optimization of each newly added Voronoi site is not sufficient, this will result in the global porosity sudden increase. Additionally, we measured the predefined particle size distribution and optimized particle size distribution over 3000 iterations in **Fig.8**. Intuitively, our optimized particle size distribution is similar to the predefined particle size distribution. Quantitatively, the error rate between optimized and predefined particle size distribution is roughly 15% according to RMSPE (Root Mean Squared Percentage Error).

Except for the distributions we used in Fig.7, we also experimented on whether it is possible to find a minor porosity under different pre-defined particle radius distributions. The top of **Fig.9** presents the multi-sized particle sampling results after 3000 iterations, while the bottom part corresponds to the pre-defined particle distribution. As a result of this experiment, setting a large

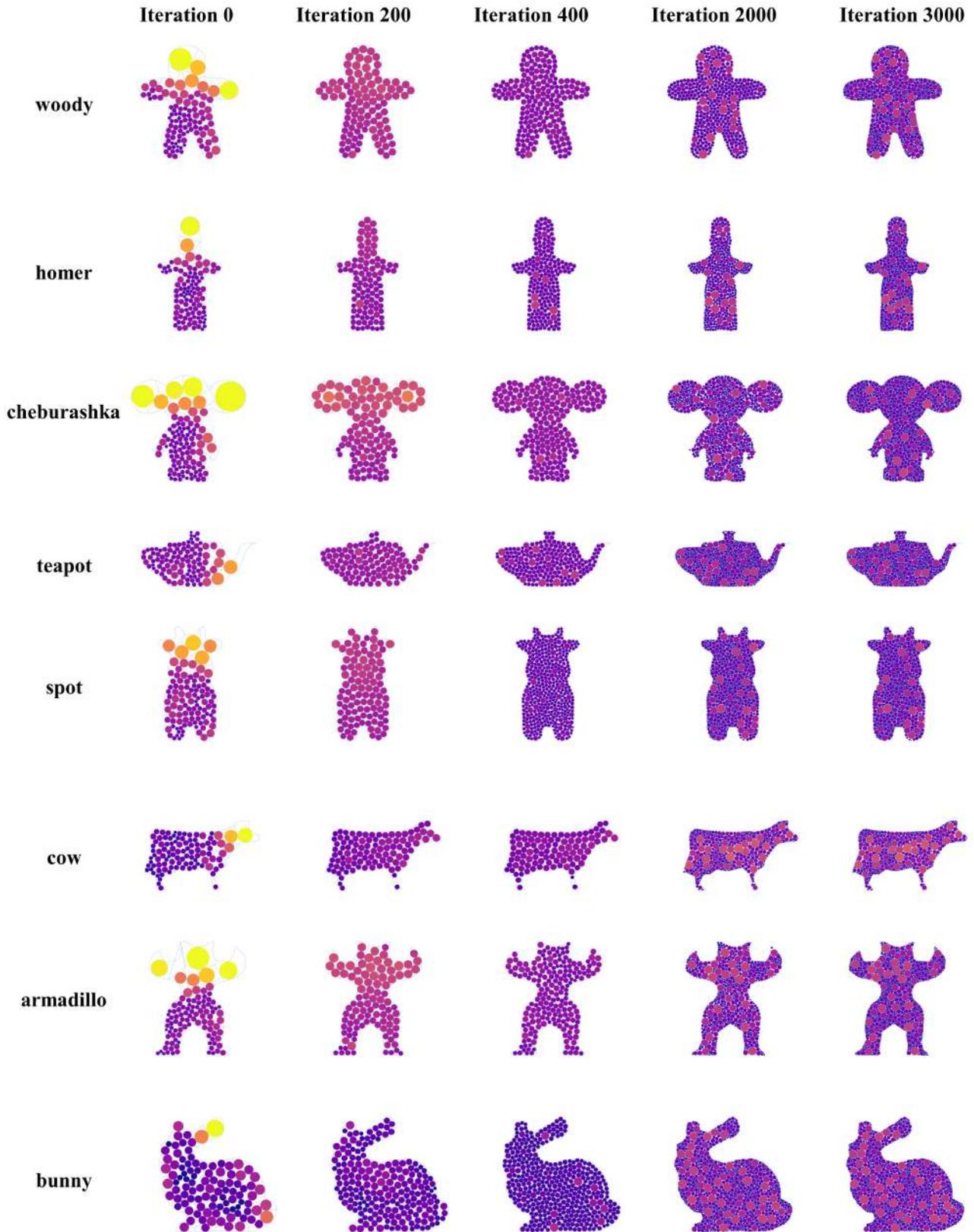


Fig. 7 Results of our multi-sized particle sampling method

number of the small size particles will be helpful to obtain a more negligible porosity.

As the right-most case in Fig.9 shows that our method allows to achieve a more complicated version of the parti-

cle distribution (containing three peaks), but in this case it will be hard to promote the packing density.

6. Discussion

Compared to uniform particle sampling method. Both standard gridding approach⁵⁾ and Poisson disk sampling⁴⁾ can keep no overlapping between particles. Nevertheless, neither of them can improve the packing density in a limited space while ensuring that the particles do not overlap with each other. In addition, as these two methods do not obtain a balanced representation at the boundary, the sampled particle clouds hardly preserve the polygon’s original shape. Figure 7 demonstrates that the particles sampled by our method can ensure a smooth boundary and that there is no overlap between particles while increasing the packing density (i.e., the global porosity is effectively reduced, as shown in Fig.5(a)). SPH-based particle sampling methods^{6),7)} can provide a uniform distribution at the boundary and produce a dense cloud of particles. Unfortunately, due to the specialty of the SPH-based method, it is not ensured that there is no overlapping between the sampled particles. It could be applied to SPH-related algorithms, but it will produce unstable simulation results when applied to DEM-related methods. Compared to the SPH-based method⁷⁾, our algorithm could provide a stable DEM simulation, as shown in Fig.10.

Compared to multi-sized particle sampling method. Most of the adaptive particle sampling methods¹⁾⁻³⁾ focused on the appearance of detailed fluid surfaces or water splashes, and they had the property in which the particle radius became larger and closer to the center. As these methods are also based on SPH, they have the same issues as SPH-based blue noise sampling⁷⁾

. Furthermore, we also provide a comparison experiment with the Poisson disk-based multi-sized particle sampling in Fig.10, which fails to achieve stable simulation results due to the overlapping particle problem.

Compared to porosity with different sampling method. While we illustrate the instability simulation when the Poisson disk sampling and SPH-based blue noise sampling methods are applied to DEM in Fig.10, the instability problem can be simply solved by detecting the overlapping parts between particles and modifying the particle radius. We admit that this approach is feasible but will induce a higher porosity, and the particle radius distribution is hard to control. Nevertheless, our approach ensures no overlap between particles and promotes the packing density (minor porosity). To prove that our method can produce a relatively more minor porosity, we adjust the size of each particle (the radius of each particle is reduced by 1/2 the amount of overlap) obtained by Poisson disk sampling and SPH-based blue noise sampling so that there is no overlap between particles. As shown in Table 1, we statistically measure the porosity of 11 different shapes of polygons respectively by using Poisson disk sampling, SPH-based blue noise sampling, and our method. Based on the results, the average porosity obtained by our sampling method is the most minor (0.165) and much better than the methods Poisson disk sampling (0.697) and SPH-based blue noise sampling (0.629).

7. Conclusion

This paper proposed to improve the packing density of particle sampling in a restricted space as well as keeping no overlap between particles. Our method allows users applying stable physics simulations based on DEM methods and related frameworks. We have introduced a multi-sized particle sampling method based on porosity optimization, which can obtain a particle distribution by minimizing the porosity in a pre-defined 2D polygon. The comparative experiments with alternative particle sampling methods demonstrate that we can deliver a stable DEM-based sand simulation.

While the experimental results have proven that our method can adapt to a variety of 2D polygons for multi-sized particle sampling, it still exists the case of undesirable sampling results as shown in Fig.11. These cases showed that if a particle with a large radius preset near a narrow area, it will make the effectiveness of local op-

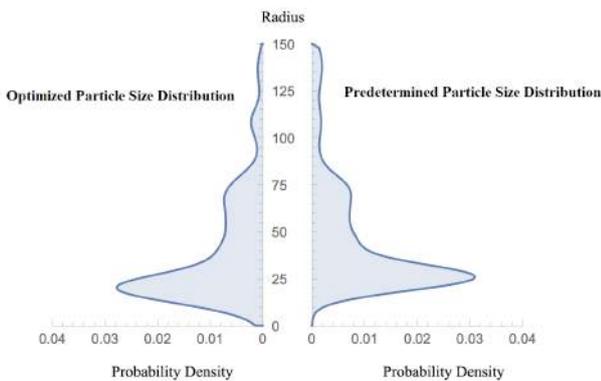


Fig. 8 Optimized and predetermined particle size distribution

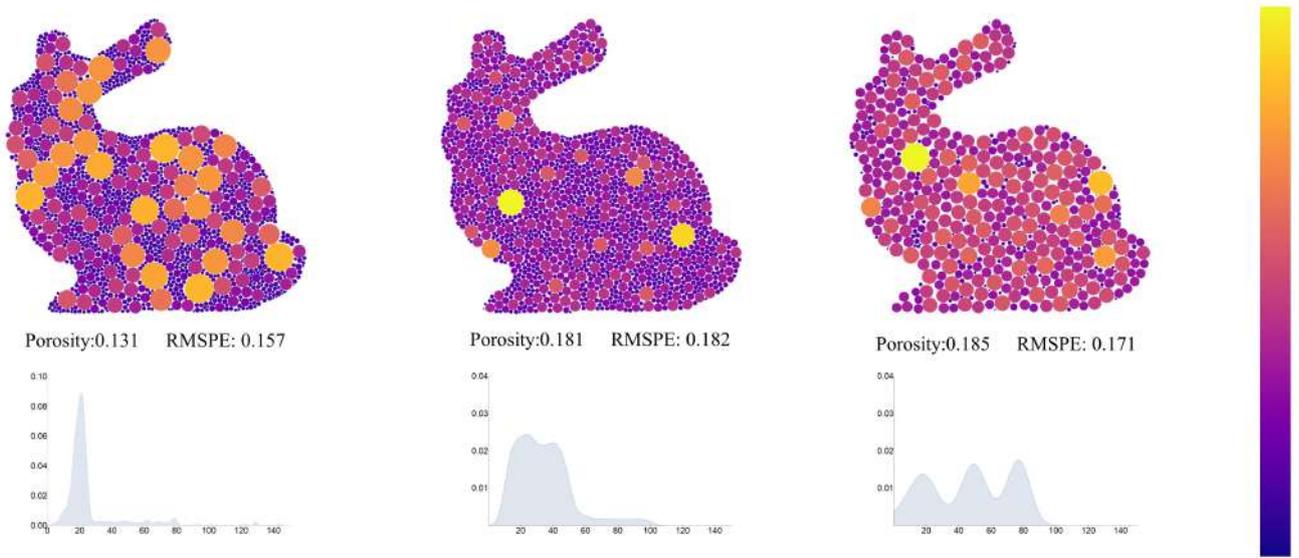


Fig. 9 Optimization results obtained from different pre-defined particle radius distributions

imization quite unfavorable. Moreover, while our approach is designed to meet the pre-defined particle radius distribution by adjusting Voronoi cells shape and size continuously, it lacks how to keep the balance of the global particle distribution, especially for the larger-sized particles. In the future, we would like to handle these problems through an extra constraint on the particle position distribution. Furthermore, we expect to explore a way to extend our method to 3D multi-sized particle sampling and adapt it to more physical simulation scenarios.

Acknowledgements

We would like to thank the anonymous reviewers for their valuable suggestions and comments. This work was supported by JSPS KAKENHI Grant Number JP20K11839 and JST SPRING Grant Number JP-MJSP2124.

References

- 1) B. Adams, M. Pauly, R. Keiser, L. J. Guibas: "Adaptively Sampled Particle Fluids", ACM Trans. on Graphics, Vol.26, No.3, pp.48:1-7 (2010).
- 2) R. Winchenbach, H. Hochstetter, A. Kolb: "Infinite Continuous Adaptivity for Incompressible SPH", ACM Trans. on Graphics, Vol.36, No.4, pp.102:1-10 (2017).
- 3) R. Winchenbach, A. Kolb: "Optimized Refinement for Spatially Adaptive SPH", ACM Trans. on Graphics, Vol.40, No.1, pp.8:1-15 (2021).
- 4) R. Bridson: "Fast Poisson Disk Sampling in Arbitrary Dimensions", Proc. of ACM SIGGRAPH Sketches, pp.22-es (2007).
- 5) M. Schwarz, H. P. Seidel: "Fast Parallel Surface and Solid Voxelization on GPUs", ACM Trans. on Graphics, Vol.29, No.6, pp.179:1-10 (2010).
- 6) H. Schechter, R. Bridson: "Ghost SPH for Animating Water", ACM Trans. on Graphics, Vol.31, No.4, pp.61:1-8 (2012).
- 7) M. Jiang, Y. Zhou, R. Wang, R. Southern, J. -J. Zhang: "Blue Noise Sampling Using an SPH-based Method", ACM Trans. on Graphics, Vol.34, No.6, pp.211:1-11 (2015).
- 8) F. De Goes, C. Wallez, J. Huang, D. Pavlov, M. Desbrun: "Power Particles: An Incompressible Fluid Solver Based on Power Diagrams", ACM Trans. on Graphics, Vol.34, No.4, pp.50:1-11 (2015).
- 9) X. Zhai, F. Hou, H. Qin, A. Hao: "Fluid Simulation with Adaptive Staggered Power Particles on GPUs", IEEE Trans. on Visualization and Computer Graphics, Vol.26, No.6, pp.2234-2246 (2018).
- 10) M. Aanjaneya, M. Gao, H. Liu, C. Batty, E. Sifakis: "Power Di-

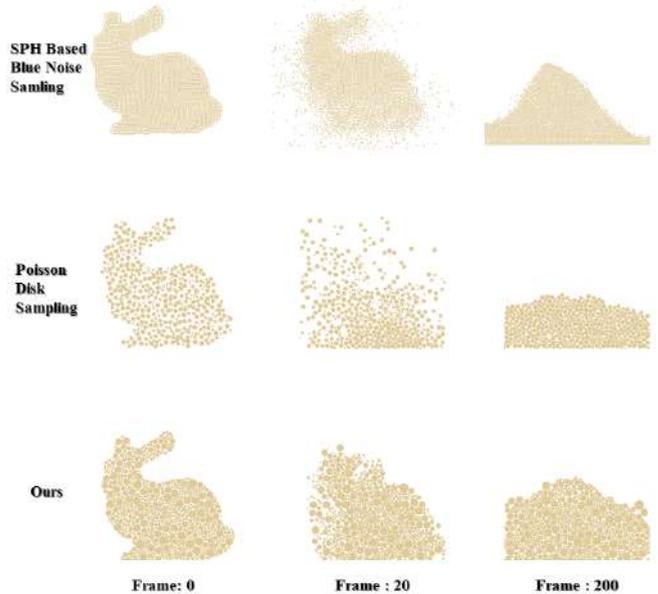
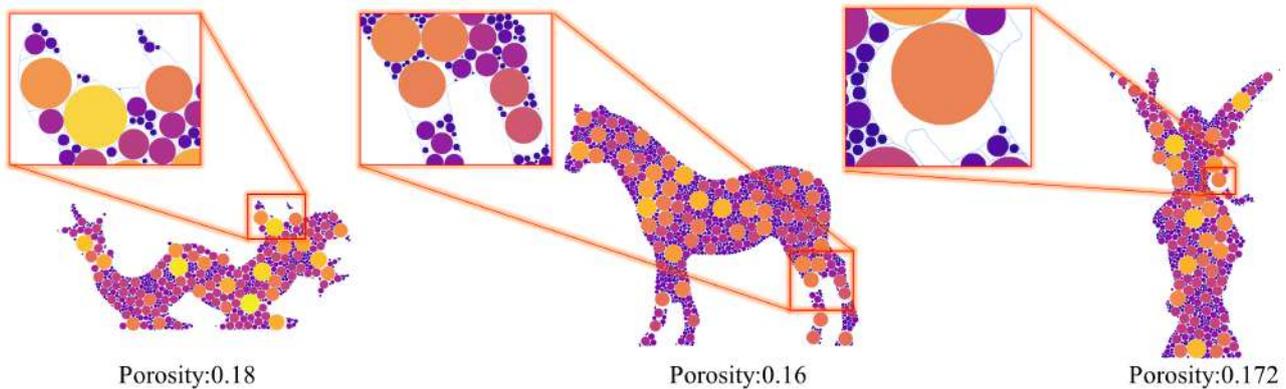


Fig. 10 Comparative experiments in sand simulation

Table 1 Table of porosity data with different particle sampling method

Example Name	Poisson Disk Sampling	SPH Based Blue Noise Sampling	Ours
Dragon	0.677	0.616	0.18
Horse	0.625	0.602	0.16
Lucy	0.605	0.613	0.172
Woody	0.703	0.598	0.161
Homer	0.708	0.655	0.172
Cheburashka	0.706	0.599	0.17
Teapot	0.713	0.635	0.164
Spot	0.772	0.653	0.168
Cow	0.704	0.649	0.17
Armadillo	0.695	0.604	0.174
Bunny	0.761	0.704	0.124
Average	0.697	0.629	0.165

**Fig. 11 Some of the failure cases**

- agrams and Sparse Paged Grids for High Resolution Adaptive Liquids”, *ACM Trans. on Graphics*, Vol.36, No.4, pp.140:1-12 (2017).
- 11) N. Ray, D. Sokolov, S. Lefebvre, B. Lévy: “Meshless Voronoi on the GPU”, *ACM Trans. on Graphics*, Vol.37, No.6, pp.265:1-12 (2018).
 - 12) X. Liu, L. Ma, J. Guo, D. -M. Yan: “Parallel Computation of 3D Clipped Voronoi Diagrams”, *IEEE Trans. on Visualization and Computer Graphics*, Vol.28, No.2, pp.1363-1372 (2020).
 - 13) J. Zheng, T. -S. Tan: “Computing Centroidal Voronoi Tessellation Using the GPU”, *Proc. of the ACM Symposium on Interactive 3D Graphics and Games*, pp.14:1-9 (2020).
 - 14) J. Basselin, L. Alonso, N. Ray, D. Sokolov, S. Lefebvre, B. Lévy: “Restricted Power Diagrams on the GPU”, *Computer Graphics Forum*, Vol.40, No.2, pp.1-12 (2021).
 - 15) M. Balzer, T. Schlömer, O. Deussen: “Capacity-Constrained Point Distributions: A Variant of Lloyd’s Method”, *ACM Trans. on Graphics*, Vol.28, No.3, pp.86:1-8 (2009).
 - 16) M. Balzer: “Capacity-Constrained Voronoi Diagrams in Continuous Spaces”, *Proc. of the IEEE Symposium on Voronoi Diagrams*, pp.79-88 (2009).
 - 17) Y. Liu, W. Wang, B. Lévy, F. Sun, D. -M. Yan, L. Lu, C. Yang: “On Centroidal Voronoi Tessellation—Energy Smoothness and Fast Computation”, *ACM Trans. on Graphics*, Vol.28, No.4, pp.101:1-17 (2009).
 - 18) R. Chen, C. Gotsman: “Parallel Blue-noise Sampling by Constrained Farthest Point Optimization”, *Computer Graphics Forum*, Vol.31, No.5, pp.1775-1785 (2012).
 - 19) M. Balzer, O. Deussen, C. Lewerentz: “Voronoi Treemaps for the Visualization of Software Metrics”, *Proc. of the ACM Symposium on Software Visualization*, pp.165-172 (2005).
 - 20) A. Noca, U. Brandes: “Computing Voronoi Treemaps: Faster, Simpler, and Resolution-independent”, *Computer Graphics Forum*, Vol.31, No.3, pp.855-864 (2012).
 - 21) W. Rungjiratananon, Z. Szego, Y. Kanamori, T. Nishita: “Real-time Animation of Sand-water Interaction”, *Computer Graphics Forum*, Vol.27, No.7, pp.1887-1893 (2008).
 - 22) T. Lenaerts, B. Adams, P. Dutré: “Porous Flow in Particle-based Fluid Simulations”, *ACM Trans. on Graphics*, Vol.27, No.3, pp.49:1-8 (2008).
 - 23) T. Lenaerts, P. Dutré: “Mixing Fluids and Granular Materials”, *Computer Graphics Forum*, Vol.28, No.2, pp.213-218 (2009).
 - 24) R. Narain, A. Golas, M. -C. Lin: “Free-flowing Granular Materials with Two-way Solid Coupling”, *ACM Trans. on Graphics*, Vol.29, No.6, pp.173:1-10 (2010).
 - 25) G. Klár, T. Gast, A. Pradhana, C. Fu, C. Schroeder, C. Jiang, J. Teran: “Drucker-prager Elastoplasticity for Sand Animation”, *ACM Trans. on Graphics*, Vol.35, No.4, pp.103:1-12 (2016).
 - 26) T. Yang, J. Chang, B. Ren, M. -C. Lin, J. -J. Zhang, S. -M. Hu: “Fast Multiple-fluid Simulation Using Helmholtz Free Energy”, *ACM Trans. on Graphics*, Vol.34, No.6, pp.201:1-11 (2015).
 - 27) T. Yang, J. Chang, M. -C. Lin, R. R. Martin, J. -J. Zhang, S. -M. Hu: “A Unified Particle System Framework for Multi-phase, Multi-material Visual Simulations”, *ACM Trans. on Graphics*, Vol.36, No.6, pp.224:1-13 (2017).
 - 28) A. P. Tampubolon, T. Gast, G. Klár, C. Fu, J. Teran, C. Jiang, K. Museth: “Multi-species Simulation of Porous Sand and Water Mixtures”, *ACM Trans. on Graphics*, Vol.36, No.4, pp.105:1-11 (2017).
 - 29) M. Gao, A. Pradhana, X. Han, Q. Guo, G. Kot, E. Sifakis,

- C. Jiang: "Animating Fluid Sediment Mixture in Particle-laden Flows", *ACM Trans. on Graphics*, Vol.37, No.4, pp.149:1-11 (2018).
- 30) Y. Yue, B. Smith, P. -Y. Chen, M. Chantharayukhonthorn, K. Kamrin, E. Grinspun: "Hybrid Grains: Adaptive Coupling of Discrete and Continuum Simulations of Granular Media", *ACM Trans. on Graphics*, Vol.37, No.6, pp.283:1-19 (2018).
- 31) X. Wang, M. Fujisawa, M. Mikawa: "Visual Simulation of Soil-Structure Destruction with Seepage Flows", *Proc. of the ACM on Computer Graphics and Interactive Techniques*, Vol.4, No.3, pp.41:1-18 (2021).
- 32) L. Zheng, Y. Yao, W. Wu, B. Xu, G. Zhang: "A Novel Computation Method of Hybrid Capacity Constrained Centroidal Power Diagram", *Computers and Graphics*, Vol.97, pp.108-116 (2021).
- 33) S. -Q. Xin, B. Lévy, Z. Chen, L. Chu, Y. Yu, C. Tu, W. Wang: "Centroidal Power Diagrams with Capacity Constraints: Computation, Applications, and Extension", *ACM Trans. on Graphics*, Vol.35, No.6, pp.244:1-12 (2016).
- 34) F. Aurenhammer: "Power Diagrams: Properties, Algorithms and Applications", *SIAM Journal on Computing*, Vol.16, No.1, pp.78-96 (1987).
- 35) J. M. Steele: *Non-Uniform Random Variate Generation*, Springer (1987)
- 36) C. B. Barber, D. P. Dobkin, H. Huhdanpaa: "The Quickhull Algorithm for Convex Hulls", *ACM Trans. on Mathematical Software*, Vol.22, No.4, pp.469-483 (1996).
- 37) P. Felkel, S. Obdrzalek: "Straight Skeleton Implementation", *Proc. of Spring Conference on Computer Graphics*, pp.210-218 (1998).

(Received March 10, 2022)
(Revised July 7, 2022)



Masahiko MIKAWA

He is currently an Associate Professor in the Faculty of Library, Information and Media Science, University of Tsukuba, Japan since 2006. He received B.Eng., M.Eng., and Ph.D. degrees from Osaka University in 1992, 1994 and 2001 respectively. He worked for NTT Access Network Systems Laboratories from 1994 to 2001, NTT Service Integration Laboratories from 2001 to 2003 and was a Lecturer in the Graduate School of Library, Information and Media Studies, University of Tsukuba from 2003 to 2006. He is a member of RSJ, SICE, SOFT and IEEE.



Xu WANG

He received his BS degree in information management and information system from Shanghai University in 2016, and MS degree in media science from Tokyo University of Technology in 2020. He is currently a PhD candidate at Physics Based Computer Graphics Laboratory, University of Tsukuba. His research interests include physics-based simulation, style transfer, and all the relevant topics in computer graphics.



Makoto FUJISAWA (*Member*)

He is currently an Associate Professor in the Faculty of Library, Information and Media Science, University of Tsukuba since 2021. He received B.Eng., M.Eng., and Ph.D. degrees in mechanical engineering from Shizuoka University in 2003, 2005, and 2008 respectively. He worked for Nara Institute of Science and Technology from 2008 to 2010 and University of Tsukuba from 2011 to 2020 as an assistant professor. His research interests include computer graphics and physics simulation. He is a member of ACM, IEEE CS, IIEEJ, IPSJ and VRSJ.

Real-Time Rendering of Oil Film with Flexible Properties

Subroto Prasetyo HUDIONO[†], Tomoya ITO^{††} (*Member*), Yuriko TAKESHIMA[†] (*Member*), Tsukasa KIKUCHI[†] (*Member*)

[†]Tokyo University of Technology, ^{††}Hachinohe Institute of Technology

<Summary> Thin-film interference is widely known for the reason of the iridescence on a bubble and spilled oil. Although it is not difficult to calculate the interference of light waves, calculating high precision interference is still considered impractical in real-time rendering. The proposed method is to separate the interference calculation by pre-computing calculation and save the value in lookup texture. But instead of using angle and thickness like other methods, we are using optical path difference as an index in lookup texture to create a flexible and compact lookup texture. The rest of the calculation, which calculates the optical path difference from the thin-film attribute will be done in real-time. Compared with other methods, the lookup texture generated using optical path difference will consume significantly less memory since it reduces the texture dimension to a one-dimensional texture. With this, we can handle different parameter changes such as refractive index and reflectance value without generating a new lookup texture or impacting the rendering performance.

Keywords: real-time rendering, thin-film interference, iridescence

1. Introduction

Iridescent surfaces are quite common in nature and has a beautiful property where the surfaces change their colors depending on the direction from which they are viewed. This occurs because light wave interference happens at the surface and causes some spectra to dim or even disappear. The most common cause is thin-film interference, a light wave interference caused by the thin film around the objects. We often see such phenomenon on soap bubbles or spilled oil, where a gradient of color appears on the surface if seen from different angles.

In computer graphics, iridescent surfaces are also often used as a material of an object, either to enhance the beauty of the objects or make them seem more realistic. It is often used as a material coating for cars or glasses. These examples can be rendered quickly since usually they have a fixed structure. Other common examples are oil spills and soap bubbles. Because of their liquid properties, the thickness of the thin film keeps changing due to surface tension. Different types of oil also have different refractive indexes, affecting the resulting color of the surface. These properties make the rendering too complex to be pre-computed while rendering it in real-time is too expensive and inefficient.

While iridescence seems to have a wide range of colors, the colors produced are due to the missing spectra after the interference. Therefore, a certain sequential pattern of color

is often seen in thin-film interference. Our approach relies on this pattern to avoid recalculating the color in every frame. By considering the physical parameters, we focus on the pre-computation of the cause of interference to make it flexible when the parameters such as refractive index, film thickness, and reflectance value change. With this, we can design an efficient look-up texture that consumes less memory with a wide range of usage.

In this paper, we will discuss the calculation that is needed to render a high accuracy interference and the importance of each thin-film parameter. From there we can separate the calculation between precomputed interference calculation and real-time attribute calculation. In addition, to produce realistic results, the rendering process from the creation of lookup texture to the real-time rendering is also explained in detail. Finally, we will compare the results in memory consumption and calculation speed to get the improvement of our method from the previous methods.

2. Related Work

Thin-film interference is often seen in glass coatings or soap bubbles. Most studies use soap bubbles to observe this behavior, and many have tried to reproduce it in computer graphics, with some excellent results attained by simulating bubble structures¹⁾ and surface tension flow on the bubble surface²⁾. However, such studies focused more on the bubble itself instead of the interference.

In the field of thin-film interference, one of the earliest

models in computer graphics was that of Smits and Meyer³⁾. Their model can accurately reproduce iridescent effects from thin films of various thicknesses. However, their method is limited to certain conditions, for example, it only applies to a smooth surface, and it does not consider inter-reflections inside the thin film. These limitations are handled splendidly in the model of Belcour and Barla⁴⁾, which uses microfacet theory for rendering the iridescent effects of varying thickness. Some works are related to iridescent color but use different physical structures. For instance, the model of Icart and Arquès⁵⁾ divides the intensity of reflected light into several intensities. A unique approach was proposed by Sun⁶⁾, whose model uses multilayer interference to simulate natural patterns found in animals and insects such as on the wings of morpho butterflies.

Despite much research on thin-film interference, there have been a few studies on its real-time rendering. One of them was an approach by Iwasaki et al.⁷⁾, which could address this problem by calculating the interference in advance and saving it as a two-dimensional texture. This pre-computation approach works well for soap bubbles where the thickness of the thin film keeps changing. However, it was inflexible since a new texture need to be rendered when there are changes in the parameter such as refractive index. Moreover, there are many duplication calculations due to using film thickness and incidence angle as a texture index. This makes the texture for interference needs too much size to render high accuracy colors.

To solve this problem, we developed a new model that can render iridescence more accurately while using less memory. Our approach focuses on creating an interference map based on the cause of the interference, which is the difference in the optical path, instead of the incidence angle and film thickness. With this approach, we can remove the influence of film properties and the angle of incoming light from the interference map. This makes it more flexible when rendering multiple objects with different film properties.

3. Physical Model of Thin Film Interference

The low surface tension of oil and its hydrophobic properties make it spread easily on objects or water surfaces under the influence of gravity. It can spread into a thin film that causes light waves to interfere with each other. This phenomenon is called thin-film interference, whereby light waves are reflected by the upper layer interfere with the diffracted light reflected by the bottom layer of the thin film, as illustrated in Fig. 1.

Since the light reflected by the bottom surface travels

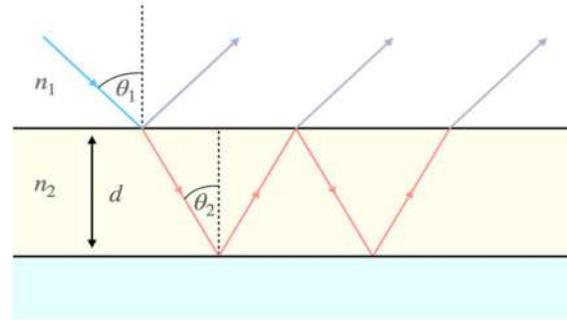


Fig.1 Light paths inside an oil film

different distances, the wavelength may be shifted, causing some frequencies to interfere and become dimmed or disappear, and leading to the reflected light missing some color spectra. This phenomenon is called destructive interference; constructive interference is when both waves having the same optical path merge and cause a brighter color.

3.1 Interference between two waves

Interference between light waves happens because the wave inside the thin film travels a different distance and causes shifting when merged with other waves on the surface of the film. This difference in distance is called optical path difference (OPD). Film thickness and light angle are the main properties that determine OPD. Depends on the direction of the light comes, the light travels farther inside the film. This is the reason why we see different colors when we see a thin-film surface from a different angle.

The refractive index is also an important parameter in thin-film interference. It is a parameter to determine the speed of light when traveling inside a substance. For example, light travels at 1.5 slower inside an oil film with a refractive index of 1.5. The difference in refractive index between substances also causes the light to change direction or refract. The relationship between incidence angle, refractive angle, and refractive index is defined in Snell's law,

$$\frac{\sin(\theta_2)}{\sin(\theta_1)} = \frac{n_1}{n_2}, \tag{1}$$

where θ_1 and n_1 are the incidence angle or angle from the light source and the refractive index of the first medium, and θ_2 and n_2 are the refractive angle and the refractive index of the film.

The Optical Path Difference or OPD is written as $m\lambda$, because we need to find the phase difference, the distance difference between wave needs to be measured with the

wavelength. With λ as the wavelength, and m represents the number of waves difference. The OPD can be found with the formula

$$2n_2d \cos(\theta_2) = m\lambda, \quad (2)$$

where d is the thickness of the thin film. From the OPD we can determine whether the interference is constructive or destructive. If m is an integer, the difference is a multiple of λ , and therefore the waves inside and outside of the film are aligned and the interference is constructive. Meanwhile, if m has a value of 0.5, the wave shifts by half of its length, and the interference is destructive, resulting in the wave disappearing.

Phase shift is also a common occurrence in thin-film interference. When light is reflected at the surface of the film, if the refractive index of the film is lower than that of the medium through which the light passes, a phase shift will occur. This results in a wave shift of 180° or half of the wavelength. This also affects OPD, which increases by 0.5λ if there is a phase shift.

With the OPD, we can calculate the intensity after interference at the surface of the thin film. We can define an equation between reflected wave and refracted wave as,

$$\begin{aligned} L(x, \lambda) &= Ar \cos(kx) + At \cos(k(x + m\lambda)), \\ &= Ar \cos(0) + At \cos(km\lambda), \\ &= Ar + At \cos(km\lambda), \end{aligned} \quad (3)$$

where A is the peak amplitude of the light wave before entering inside the film, r is the reflectance, t is the transmittance value, k is the wavenumber, x is the initial phase of the wave, and $m\lambda$ is the OPD. Since we use the difference in the optical path, interference can be assumed at the start of angular frequency which makes x to be 0.

In the actual case, the light that enters inside the film is reflected and attenuated several times until losing its energy. Every time a portion of the light travels out of the film, it has traveled a certain distance that depends on how many times it was reflected inside the film. Thus, interference happens multiple times at different phases as illustrated in Fig. 1. With the assumption that the thin film is on a solid surface that fully reflects the light, we can define it as the equation,

$$\begin{aligned} L(\lambda) &= Ar + At^2 \cos(km\lambda) + Art^2 \cos(2km\lambda) + \\ &\quad Ar^2t^2 \cos(3km\lambda) + \dots, \\ &= Ar + At^2 \cos\left(\frac{2\pi}{\lambda}m\lambda\right) + Art^2 \cos\left(2\frac{2\pi}{\lambda}m\lambda\right) + \\ &\quad Ar^2t^2 \cos\left(3\frac{2\pi}{\lambda}m\lambda\right) + \dots \\ &= A \left[r + \sum_{n=1}^{+\infty} r^{n-1} t^2 \cos(2\pi nm) \right] \end{aligned} \quad (4)$$

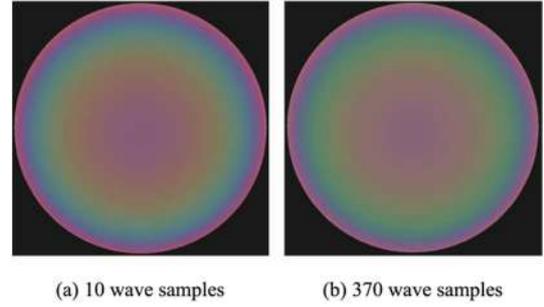


Fig.2 A difference in saturation between rendering the interference with 10 wave samples (a) and 370 wave samples (b)

where the Σ represents the total radiance from all interfered waves and n represents the number of lights that went outside the film. Since the material color is defined by albedo texture, the reflectance value of the material assumed to be 1 and can be removed from the equation.

3.2 Visible spectrum

The visible light spectrum covers wavelengths ranging from about 380 to 750 nanometers. Each wavelength represents a different color, while white light is a combination of multiple lights of different wavelengths. When the light gets inside the thin film, it contains a lot of different waves at different wavelengths. Therefore, we can say that the interference happens in all the visible light spectra that got inside the thin film.

We can accurately calculate the output color by determining the radiance of each light that enters inside the thin film. This can be defined as,

$$L = \int L(\lambda) d\lambda \quad (5)$$

where $L(\lambda)$ is the spectral radiance of the light at a certain wavelength. The number of samples that we use in this interference equation also affects the accuracy of the final color. As shown in **Fig. 2**, fewer samples tend to result in over-saturated color output, thus a higher number of wave samples is required to accurately render the color when calculating the total radiance.

4. Rendering Method

Simulating a large number of wavelength samples to produce realistic output colors in thin-film interference is computationally expensive. Moreover, rendering sRGB color space using wavelengths with high accuracy could be impractical in real-time rendering. Especially when there are variations in viewing angle and thin-film properties, the

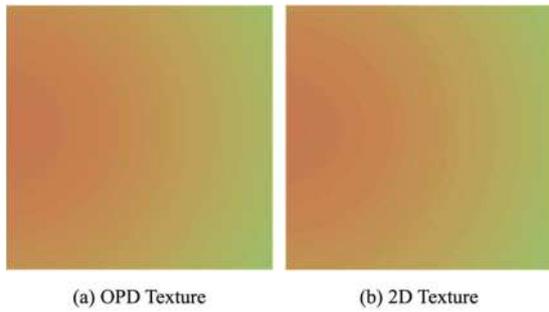


Fig.3 An OPD texture (a) takes 640 bytes and a 2D texture (b) takes 4,096 bytes of memory

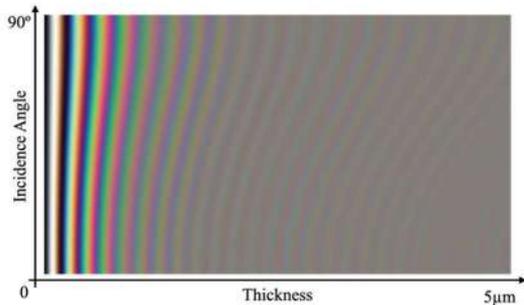


Fig.4 Two-dimensional interference map from thicknesses and incidence angles without considering the reflectance value

output color will have to be recomputed⁴).

To date, the simplest way to render thin-film interference in real-time rendering has been to pre-compute all possible angles and thicknesses from thin-film interference use them as a texture⁷). Although this could be used in real-time rendering, it is not flexible enough to handle the change in parameters such as refractive index and reflectance value. Moreover, it contains duplicate values along the Y-axis and uses memory consumption ineffectively. As seen in **Fig. 3**, 2D texture can consume more memory with lower accuracy compared to using OPD texture.

4.1 Optical path difference

We propose a new rendering method in which, instead of using thickness and angle to determine the resulting color in a two-dimensional texture, we map the output color in a one-dimensional texture and use OPD as an index.

The main cause of differences in color output in thin-film interference is the shift in wavelength due to the different distances traveled, or OPD. First, the incidence angle and film thickness are input to determine the OPD, however, using both of them as an index is inefficient since

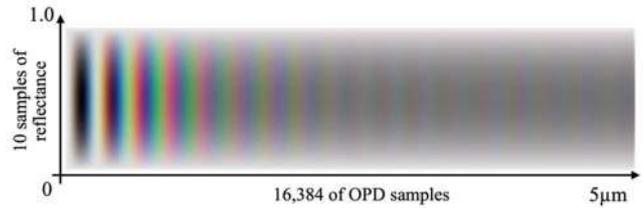


Fig.5 Lookup texture by calculating the interference with every possible OPD and reflectance value

the texture will save multiple duplicate values. This can be seen in **Fig. 4**. By ignoring the reflectance value to make the pattern visible, the value is mostly the same between the Y axis, only stretched because of the incidence angle.

In generating an interference map, we calculate interference with the formula from Eq. (4) to simulate the interference that happens between the light inside the film. This interference between light waves has less effect with longer OPD, as we can see in Fig. 4. From this, we can simulate the interference that happens with OPD below 5 μ m.

The input for the calculation will be the OPD $m\lambda$ and the wavelength (λ). To simulate the interference accurately, we sample 370 wavelengths from 350 to 750. With the same OPD on every wavelength, we will get different multiplication numbers (m) that determine whether there is destructive interference or not. Finally, we average the results after changing the wavelength to RGB colors and save it in the texture.

4.2 Reflectance value

Another factor that can affect interference is the reflectance value of the thin film. Reflectance value is the proportion of the wave when separated into a refracted wave and a reflected wave when hitting the thin-film surface. Since the reflectance value is the proportion of the divided light, we can represent it as a scale number from 0 to 1.

With reflectance value, the lookup texture will become a 2-dimensional texture with reflectance value as the Y index. The result of this texture can be seen in **Fig.5**. We will have a bigger lookup texture, but since the reflectance value is linear and doesn't need high precision, we can use a small value with the incremental of 0.1 precision. Moreover, the result is horizontally symmetrical, so we only need to calculate half of the value which is 0 to 0.5. This will increase the lookup texture by 6 times larger, but still much smaller by using a fully 2-dimensional texture.

4.3 Wavelength to RGB color space

The output of interference calculation (Eq. (4)) is the radiance of each wavelength, from this we use CIE's color matching functions to transform the wavelength (λ) into RGB⁸). This color-matching function is based on the equation,

$$\begin{aligned} X &= \int_{\lambda} L(\lambda)\bar{x}(\lambda)d\lambda, \\ Y &= \int_{\lambda} L(\lambda)\bar{y}(\lambda)d\lambda, \\ Z &= \int_{\lambda} L(\lambda)\bar{z}(\lambda)d\lambda, \end{aligned} \quad (7)$$

Given $L(\lambda)$ is the spectral radiance of the wavelength λ . This tristimulus value (XYZ) can be transformed to RGB using the following matrix⁹,

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.24048 & -1.53715 & -0.49854 \\ -0.96926 & 1.87599 & 0.04156 \\ 0.05565 & -0.20404 & 1.05731 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (8)$$

Then the RGB value will be scaled by the number of wavelength samples used to simulate the interference.

5. Results

The accuracy of our approach was tested by comparing the result of interference color with a photograph of a real soap film from Atkin and Eliot¹⁰ (Fig. 6). With color temperature adjustment, our simulated oil film, rendered in a linear incremental of film thickness, could well reproduce the color pattern of the photograph of interference. Since our approach is using simulated optical path difference to calculate the interference, the results would be the same even for materials with different refractive indexes.

We have implemented our approach in GLSL shaders using our custom renderer running in Vulkan with AMD Radeon Pro 450. We tested the speed performance of calculating OPD in real-time rendering with physically based rendering (PBR) on a metallic surface object. The rendering times for each frame are given in Table 1.

The first row shows the rendering times without interference for a normal object and material by PBR. The second row shows the rendering time when the interference

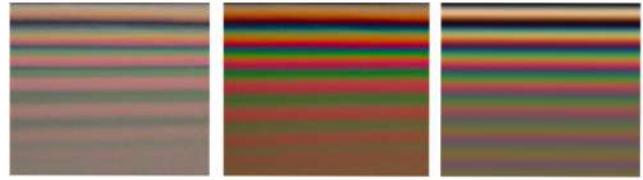


Fig.6 Results of thin-film interference photographed for a real soap film (left¹⁰ and middle¹) and our simulated oil film (right)

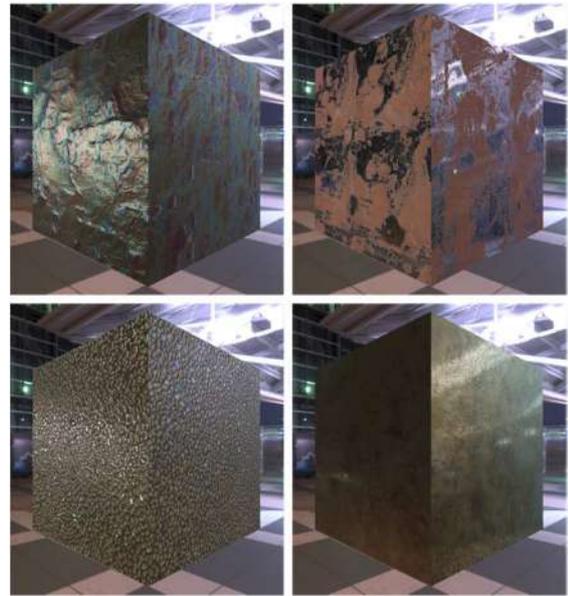


Fig.7 Thin-film interference implemented in a cube with different materials

was calculated directly while rendering the scene. The last two rows show the results of testing the scene with 2D format texture and our approach. From this table, we know that a naive approach to calculating the thin-film interference with high accuracy isn't feasible in real-time rendering. And while our approach takes more rendering time the difference is too small and didn't scale by the complexity of the scene

The memory consumption of OPD texture is also significantly smaller compared with 2D texture lookup. Because of this, we can increase the resolution without worrying about the size of the texture lookup. We calculate the estimated size of the texture in Table 2

Table 1 Rendering time per frame in millisecond

	Sphere	Bunny	Texture	Bubble
No Interference	8.8 ms	12.8 ms	28.0 ms	15.2 ms
Naïve	134.6 ms	211.8 ms	114.9 ms	147.8 ms
Texture (2D)	9.6 ms	13.4 ms	28.3 ms	16.8 ms
Ours (1D)	10.1 ms	13.6 ms	28.3 ms	17.0 ms

Table 2 Memory consumption from different sample size

Resolution	2D Texture	OPD Texture	OPD with Reflective
1024 samples	4 MB	4 KB	24 KB
2048 samples	8 MB	8 KB	48 KB
8192 samples	32 MB	32 KB	192 KB
16384 samples	64 MB	64 KB	384 KB



Fig.8 Rendered soap bubble with thin-film interference

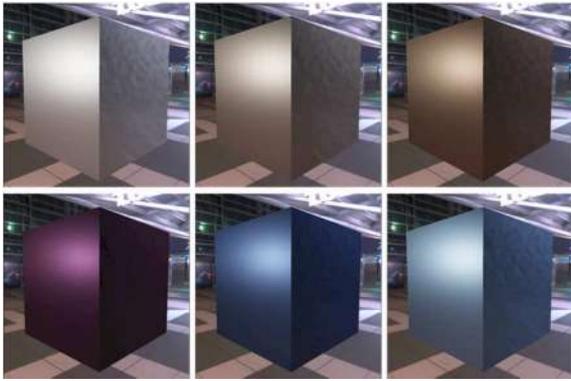


Fig.9 Tempered steel color reproduced by increasing the film thickness to simulate the oxidation

In 2D texture from Iwasaki et al. ⁷⁾, the texture is generated from the film's thickness and incidence angle. Since both are important factors for color accuracy, they need to have high accuracy for generating a lookup texture. With this, we generate the lookup texture with at least 1024 samples for the incidence angle with accuracy per 0.1 degree.

We tested our method with physically based rendering on different materials to test the color accuracy (**Fig. 7**). While the surface that has a thin film will reflect light, some parts of the material have low reflectance, and might not show any color differences. With that in mind, we tested our approach by implementing the interference on the part of the object which reflects light.

We also reproduced objects that have a natural thin-film interference property like soap bubble (**Fig. 8**) and tempered steel (**Fig. 9**). The heating process in tempered steel can cause a layer of iron oxide to form on the surface of the steel. As the temperature of the steel increases, the thickness of the

iron oxide will also increase, forming a thin film and causing the steel color to change. Therefore, we tried to reproduce the thin-film interference on the tempered steel by increasing the thickness of the film from $0.12\mu\text{m}$ to $0.24\mu\text{m}$.

To test the importance of refractive index as a parameter of thin films, we tried to compare the difference when increasing the refractive index and the film thickness in a sphere. From the test in **Fig.10**, we know that the change in film thickness has a different effect from the change in refractive index.

Increasing the thickness of the film will give more gradient and make the material more sensitive to color change, this is because by increasing the thickness, the OPD will also increase and ended up using a wider area in the lookup texture. Meanwhile, the change in the refractive index gives us shorter gradient colors. This is because the refractive index also affects the refractive angle and causes the view angle to have a smaller effect on the OPD.

One of the examples of an object with a high refractive index is tempered steel. Tempered still tends to have the same color even though viewed from different angles. By reducing the refractive index, the area of the lookup table that is used to get the color is also scaled down. This result cannot be achieved when using a 2D lookup texture based on film thickness and incidence angle.

6. Conclusion and Future Works

We have presented a new approach to generate lookup texture with optical path difference as an index. One of the advantages of our method is the flexible properties of a thin film that can be changed without generating a new lookup texture. This includes the film thickness, refractive index, and reflective value, while the texture used from other methods is only able to change the film thickness.

Moreover, with the pre-computation of the interference, there will be no limitation in the first rendering time, and a more accurate RGB model could be used to achieve more accurate color when transforming wavelength into color without affecting the rendering time. This could be useful when rendering spilled oil on a road or car material in games where the scene needs to have a high frame rate.

Our contribution mainly focuses on providing a new flexible way to apply iridescent materials to different types of objects with high accuracy in real-time rendering. As a result, a scene with different types of objects with different material properties could be rendered accurately by only using one small interference texture in real-time. It is also

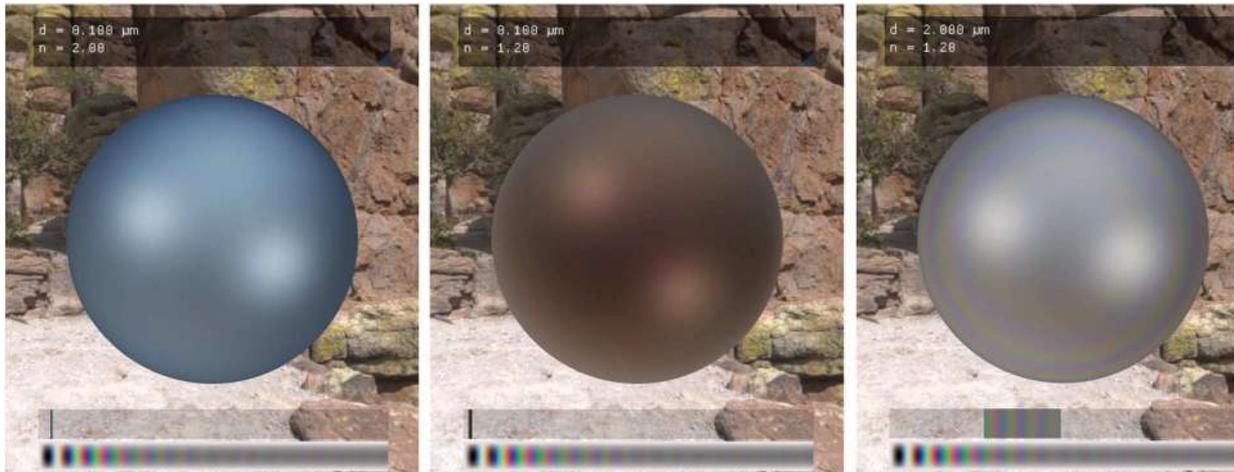


Fig.10 The difference in changing the thickness and refractive index of the thin film. The left image is the sphere with a high refractive index, the middle image is the default parameters, and the right image is the image with a thicker thin film

possible to implement the proposed method directly into production without any adjustments.

The limitation of our approach is the assumption that the object only has a single film and a smooth surface in which the surface will fully reflect the light. When the object has multiple films, the interference becomes much more complex that the interference map might not be possible to produce an accurate result. One of the reliable methods is to render all the interference in real-time but it is not possible for real-time rendering. But it is rare to find an object with more than one thin film.

Future work includes rendering multiple layers of thin films. The proposed model focused mainly on rendering one film on the object's surface. Meanwhile, in a spilled oil, the light that passed through also interact with the water below the oil film making the object has two films with different refractive index and different thickness. This could be an interesting topic to study considering spilled oil is a common occurrence in the real world.

Finally, even though real-time rendering techniques have improved greatly together with the development of computer hardware in recent years, providing a real-time rendering with high realism is still a challenge. A lot of aspects need to be considered, and even small improvements could help to save up time during rendering. In the case of thin-film interference, the color outputs already have great accuracy, so instead of the output color, a film behavior like how the thickness move in oil film due to the surface tension is also an interesting part of future work.

References

- 1) S. Ishida, P. Synak, F. Narita, T. Hachisuka, C. Wojtan: "A Model for Soap Film Dynamics with Evolving Thickness," *ACM Trans. on Graphics*, Vol.39, No.4, pp. 31:1–31:11 (2020).
- 2) W. Huang, J. Iseringhausen, T. Kneiphof, Z. Qu, C. Jiang, M. B. Hullin: "Chemomechanical Simulation of Soap Film Flow on Spherical Bubbles," *ACM Trans. on Graphics*, Vol.39, No.4 (2020).
- 3) B. Smits, G. Meyer: "Newton's Colors: Simulating Interference Phenomena in Realistic Image Synthesis," (1992).
- 4) L. Belcour, P. Barla: "A Practical Extension to Microfacet Theory for the Modeling of Varying Iridescence," *ACM Trans. on Graphics*, Vol.36, No.4, p. 14 (2017).
- 5) I. Icart, D. Arques: "An Illumination Model for a System of Isotropic Substrate- Isotropic Thin Film with Identical Rough Boundaries," In *Proceedings of the 10th Eurographics conference on Rendering*, Eurographics Association, Goslar, DEU, 261–272 (1999).
- 6) Yinlong Sun: "Rendering Biological Iridescences with RGB-Based Renderers," *ACM Trans. on Graphics*, Vol.25, No.1, pp. 100-129 (2006).
- 7) K. Iwasaki, K. Matsuzawa, T. Nishita: "Real-time rendering of soap bubbles taking into account light interference," *Proceedings Computer Graphics International*, pp. 344-348 (2004).
- 8) C. Wyman, Peter-Pike J. Sloan, P. Shirley: "Simple Analytic Approximations to the CIE XYZ Color Matching Functions," *Journal of Computer Graphics Techniques*, Vol.2, No.2 (2013).
- 9) R. Durikovic, Ryou Kimura: "GPU Rendering of the Thin Film on Paints with Full Spectrum," *Tenth International Conference on Information Visualisation (IV'06)*, pp. 751-756 (2006).
- 10) L. J. Atkins, R. C. Elliott: "Investigating thin film interference with a digital camera," *American Journal of Physics*, Vol.78 (2010).

(Received March 11, 2022)



Subroto Prasetyo HUDIONO

He received his B. Sc. degree from Institut Sains dan Teknologi Terpadu Surabaya in 2014. After entered Tokyo University of Technology, Department of Media Science in 2020, he completed his master's degree program in 2022. His research interests include computer graphics, game, and real-time rendering.



Tomoya ITO *(Member)*

He received the Ph.D degree from Iwate University, Japan in Engineering. He has been a Professor in the Department of Engineering, at Hachinohe Institute of Technology since 2004. His research interests include computer graphics, game, and media art. He is a member of Information Processing Society of Japan, Game Amusement Society, and the Society of Art and Science.



Yuriko TAKESHIMA *(Member)*

She received her B. Sc., M. Sc., and Ph.D. degrees in Computer Sciences from Ochanomizu University in 1994, 1996, and 1999, respectively. She is currently a Professor at School of Media Science, Tokyo University of Technology. Her research interests include volume visualization, topology-based visualization, and visualization environments.



Tsukasa KIKUCHI *(Member)*

He received his Ph.D degree from Iwate University in Engineering in 1999. He is currently a Professor at School of Media Science, Tokyo University of Technology. His research interests include computer graphics, visual simulation, procedural animation and contents design. He is a member of Information Processing Society of Japan, and the Society of Art and Science.

Eyeliner Looked “Attractive” by Eyelids Shape Using Eyeliner Models

Rena OKURI[†], Shuhei KODAMA[†], Tokiichiro TAKAHASHI^{††} (*Fellow*)

[†]Tokyo Denki University, ^{††}ASTRODESIGN,Inc.

<Summary> “Eyeliner,” a makeup method makes the eyes look larger, and makes the face so-called ‘attractive.’ Eyelids shapes of Asian women are roughly categorized into “double,” “hooded,” and “single.” We assumed that the most attractive eyeliner for her would depend on her eyelids shape. In this paper, we quantitatively measured the length and the thickness of eyeliner (eyeliner style) that make the eyes appear the largest without any sense of incongruity for each eyelids shape and investigated whether the eyeliner style is attractive. The purpose of this paper is to clarify attractive eyeliner styles by eyelids shape. The measurement was conducted using digital makeup with eyeliner models devised in this study. The results showed that the eyeliner that makes the eyes look the largest without any sense of incongruity is the most attractive eyeliner for each eyelids shape.

Keywords: makeup, eyeliner, eyeliner model

1. Introduction

Facial impression is one of the important factors in impressing personal so-called ‘attractiveness.’ According to Dipboye et al.¹⁾, an attractive person even has advantages in social life. Facial features are related to facial attractiveness and faces with features of high symmetry and averageness are judged to be attractive²⁾. Furthermore, according to experimental studies on facial morphology, large eyes have one of the features that increases the attractiveness of a woman face²⁾ because they give the impression to be feminine and youthful^{3),4)}.

It is difficult to change the shape of the face itself to have these features. However, makeup, applying various cosmetics to the face, is a way to approach the face with attractive features only in appearance without changing the shape of the face. Makeup has the effect of increasing facial attractiveness and beauty^{5)–7)}.

There have been developed various makeup methods, and examining their effectiveness is an important role to improve the attractiveness with the recent growing interest in facial beauty and attractiveness. Previous work on the effects of makeup have shown that eye makeup is particularly effective in increasing attractiveness by changing the apparent features of the eyes⁸⁾.

Eye makeup has the effect of making the eyes appear larger and emphasizing them so that it can improve the face attractive facial impression. Killian et al.⁷⁾ inves-

tigated the change in attractiveness and symmetry, and then they clarified the correlation between the two by applying makeup to specific areas of the face. The results showed that the upper eyelids makeup significantly increased attractiveness. Makeup on the areas of increased attractiveness slightly decreased asymmetry. Thus, they mentioned that makeup may increase attractiveness at least partially by decreasing asymmetry. Jones et al.⁸⁾ verified the hypothesis that makeup makes facial features (eyebrows, eyes, nose, and mouth) appear larger. The results showed that both self-applied and professionally applied makeup was perceived to have enlarged the eyes. Morikawa et al.⁹⁾ revealed through experiments using psychophysical methods that the eyes with eye shadow appeared larger. They showed that the perception of eye size with eye shadow was related to both the assimilation effect of the eye and the eye shadow, and the assimilation effect of the eyebrow and the eye caused by the eye shadow filling the area between the eye and the eyebrow. These results support that changing the size of facial features with makeup increases facial attractiveness.

Moreover, eyeliner which is a makeup technique drawing a line around the eyelash line has the effect of making the eyes appear larger. Matsushita et al.¹⁰⁾ indicated a psychological explanation for this effect. They experimentally revealed that the eyeliner had an illusion effect of making the eyes appear larger based on the Delboeuf illusion, and that the color and the range of eyeliner (“up-

per eyelid only” or “upper eyelid + lower eyelid”) affect how much the eyes appear to be larger (called the amount of illusion). The Delboeuf illusion is the illusion that the inner circle of a double circle appears to be larger than that of a single circle (the inner circle is the same size as the single circle), as shown in **Fig. 1**. This illusion is caused by the assimilation of the outer and inner circles of the double circles. Okuri et al.¹¹⁾ considered that the thickness of eyeliner as well as the color and the range of eyeliner which Matsushita et al.¹⁰⁾ focused on affects the amount of illusion, so they measured the amount of illusion between the thickness of eyeliner and the size of the double eyelids using a psychophysical measurement method. Also, they conducted a questionnaire to investigate the sense of incongruity for the eyes with eyeliner of various thicknesses to determine the appropriate thickness of eyeliner for eye makeup. However, their study only examined the case of double eyelids, so they have no findings in the case of different eyelids shapes.

Eyelids shapes of Asian women are classified into single, hooded and double, and the attractiveness differs depending on the eyelids shapes. East and Southeast Asian women often have single eyelids or thin double eyelids congenitally. Women with double eyelids are sometimes said to be more attractive than single eyelids because their eyes appear larger. Jung et al.¹²⁾ measured the physical width and height of the eyes in Korean subjects with and without double eyelids to investigate whether the eyelids shapes actually make a difference in eye size. Their measurement results revealed that the width and height of the eyes were significantly longer in the double eyelids than in the single eyelids. There were a couple of reports that the effect of eye makeup differed depending on the shape of eyelids. Takehara et al.¹³⁾ examined the physical and mental attractiveness of woman’s faces with single/double eyelids and with/without eye bag makeup. Their results indicated that the faces with double eyelids were more attractive than single eyelids. The faces with eye bag makeup were generally perceived as less attractive than those without them. Nishino et al.¹⁴⁾ examined visual effects of double eyelids with psychological experimental methods and compared with the effects by eyeliner. They showed that the eye enlargement effects of double eyelids and eyeliner have additive effects. The experiment suggested that double eyelids are more attractive than single eyelids, and this effect is greater with eyeliner makeup.

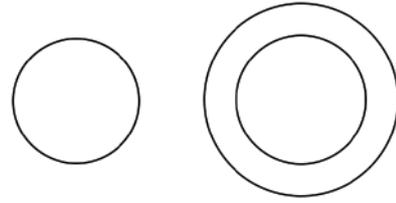


Fig. 1 An example of the Delboeuf illusion

We have briefly reviewed the studies on attractive makeup. To our knowledge, there have been no related studies that investigated the relations the length and the thickness of the attractive eyeliner by eyelids shape (hereafter, referred to as “eyeliner style”). In this paper, we quantitatively measure the length and the thickness of eyeliner that makes the eyes appear the largest without any sense of incongruity for each eyelids shape (single eyelids, hooded eyelids, and double eyelids) by continuously changing the numerical values of both the length and the thickness of the eyeliner. We numerically clarify which length and thickness of the eyeliner make the eyes look the largest without any sense of incongruity is an attractive eyeliner style since larger eyes are said to be more attractive.

2. Method

We conducted an experiment to quantitatively verify the eyeliner style that makes the eyes perceived to be the largest without any sense of incongruity and considered whether the eyeliner style has the effect of increasing attractiveness.

2.1 Participants

Thirty-five people (mean age 22.5 years, SD = 1.79, 22 participants with experienced makeup and 13 participants with inexperienced makeup) with normal visual acuity and color vision participated in the experiment. All experienced participants in this experiment were female, and all inexperienced participants were male. We explained the contents of our experiment in verbal and written, then participants consented to the experiment of their own free will. Additionally, we used statistically processed data for analysis so that individuals cannot be identified.

2.2 Environment

During the measurement, we instructed that cold lighting should be used, the brightness should be constant, and the direction of the lighting should be fixed. The display was s2430L dell monitor (24 inches). The position

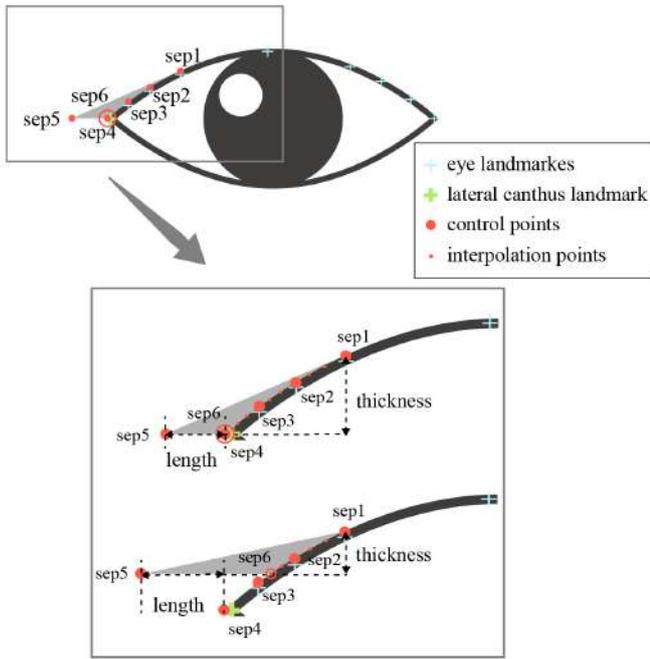


Fig. 2 A model of a single eyelids eyeliner

and orientation of the display should be fixed, and the luminance should be measured with a color luminance meter (CS-100A, Konica Minolta) to be approximately 250 cd/m². The aspect ratio of the display was 16:9. The height of the display screen was about 30.5 cm, and the observation distance was set to about 90 cm (three times the height of the display screen). All participants followed the experimental procedures.

2.3 Eyeliner models

In order to represent eyeliner digitally, we define eyeliner models based on general shape of eyeliner. The eyeliner models were determined based on the coordinates of the feature points on the upper eyelids. We use MediaPipe Face Mesh¹⁵⁾ to get the feature points. MediaPipe Face Mesh is a library provided by Google that estimates 468 3D facial landmarks in real-time. A model of the eyeliner shape for each eyelids shape is shown as follows. The models parameterized *thickness* and *length*.

Single eyelids have a structure in which the skin is not connected to the muscles that raise the eyelids, so that the skin over the eyelids is not pulled up and covers the eyelids when the eyes are open. Hooded eyelids have a similar structure on the inner side of the eye. For this reason, eyeliner models were created for the single eyelids and the hooded eyelids as far as they are visible when the eyes are open.

2.3.1 Single eyelids eyeliner

The model of a single eyelids eyeliner is shown in **Fig. 2** and it is determined by the following steps.

1. Determine *sep1*, *sep2*, *sep3*, and *sep4* from the eye landmarks.
2. Calculate a sequence of points following the eye shape by quadratic B-spline interpolation with *sep1*, *sep2*, *sep3*, and *sep4* as control points. Note that the vertical coordinates of the point sequence should be different by 1px.
3. Let *sep6* be a point in the point sequence calculated in Step 2 that matches a vertical coordinate shifted vertically from *sep1* by *thickness*.
4. The horizontal coordinate of *sep5* is the point shifted horizontally from the lateral canthus landmark by *length*, and the vertical coordinate is the point shifted vertically from *sep1* by *thickness* (same vertical coordinate as *sep6*).
5. Create a polygon with *sep1*, *sep5*, *sep6*, and interpolation points between *sep6* and *sep1*.
6. Fill in the polygon in Step 5.

2.3.2 Hooded eyelids eyeliner

The model of a hooded eyelids eyeliner is shown in **Fig. 3** and it is determined by the following steps.

1. Determine *hep1*, *hep2*, *hep3*, and *hep4* from the eye landmarks.
2. The point that is shifted *length* parallel horizontally from the lateral canthus landmark is *hep5*.
3. *hep8* and *hep7* are the same points as *hep2* and *hep3*, respectively.
4. Let *hep6* be a point that is shifted vertically from *hep4* by *thickness*.
5. Create a polygon by quadratic B-spline interpolation with *hep1*, *hep2*, *hep3*, *hep4*, *hep5*, *hep6*, *hep7*, and *hep8* as control points.
6. Fill in the polygon in Step 5.

2.3.3 Double eyelids eyeliner

The model of a double eyelids eyeliner is shown in **Fig. 4** and it is determined by the following steps.

1. Determine *dep1*, *dep2*, *dep3*, and *dep4* from the eye landmarks.
2. The point that is shifted *length* parallel horizontally from the lateral canthus landmark is *dep5*.
3. Let *dep8* be a point that is shifted vertically from *dep2* by *thickness*. Similarly, determine *dep7* from *dep3*, and *dep6* from *dep4*.
4. Create a polygon by quadratic B-spline interpolation with *dep1*, *dep2*, *dep3*, *dep4*, *dep5*, *dep6*, *dep7*, and

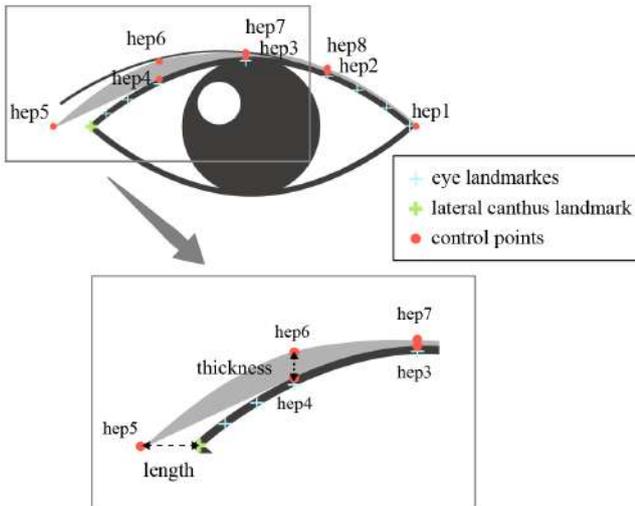


Fig. 3 A model of a hooded eyelids eyeliner

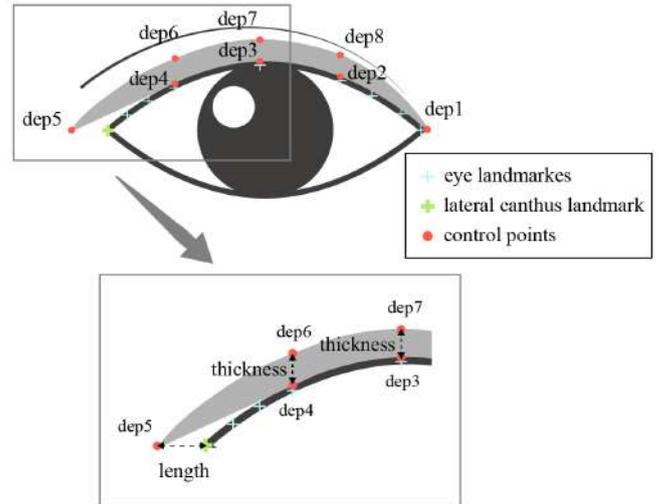


Fig. 4 A model of a double eyelids eyeliner

dep8 as control points.

5. Fill in the polygon in Step 4.

2.4 System

Figure 5(a) shows the image used in the experiment, a 669 px × 946 px face image of a Japanese woman with double eyelids without makeup. The face image was arranged with hair style that covered the eyebrows, because the eyebrows and hair may cause bias in the size perception. The right eye was created by inverting the left eye then the eyes were placed symmetrically in a sagittal line. We retouched the boundary of the right eye and the facial skin by Photoshop to make them natural. We created two other face images with single eyelids and hooded eyelids by retouching the face image with double eyelids as shown in Fig. 5(b)(c). Since the eye size of single eyelids is physically shorter than double eyelids in both width and height¹²⁾, the face image with single eyelids was created to have smaller eyes than the face image with double eyelids. The hooded eyelids were created by erasing the double eyelids line from the inner corner to the center of the eye without changing the size of the double eyelids. In this way, three models, a double eyelids model, a hooded eyelids model, and a single eyelids model were prepared.

The actual image on the screen was set to 1 px = 0.3 mm (rounded to the second decimal place) at a resolution of 1,920 px × 1,080 px. The eye width on the screen was adjusted to be approximately 2.7 cm and the eye height was approximately 1.0 cm in the hooded eyelids model and the double eyelids model based on the average interocular breadth of 3.7 cm and biectocanthion breadth of 9.0 cm, which are typical for Japanese young adult female¹⁶⁾. For the single eyelids model, the width was

approximately 2.5 cm and the height was approximately 0.8 cm.

In this experiment, we prepared a system that can change the parameters of length and thickness of the eyeliner drawn by computer graphics simulation to quantitatively verify the eyeliner style. The system allows the participants to adjust length and thickness by using the slider interface. The adjustable range is 0 mm to about 11.1 mm (0 px to 40 px) for length of three types of eyelids shapes models, 0 mm to about 4.2 mm (0 px to 15 px) for thickness of the single eyelids model, and 0 mm to about 5.5 mm (0 px to 20 px) for thickness of the hooded eyelids model and the double eyelids model. The eyeliner was drawn for each eyelids shape using the eyeliner models defined in Section 2.3.

Some of the eyeliner for each eyelids shape model is shown in **Table 1** for length manipulation and in **Table 2** for thickness manipulation. In these figures, only the left eye is cropped. The images displayed in the experiment were the entire face.

2.5 Procedure

Participants looked at the images presented by the experimental system and adjusted length and thickness of the eyeliner to get the most appropriate eyeliner for each condition.

The experimental procedure was as follows.

First, the screen showed an image of a face whose the length and the thickness of eyeliner can be changed sequentially by the participant's manipulation.

Then, the participants were asked to respond to three conditions per model (the single eyelids model, the hooded eyelids model, and the double eyelids model) by

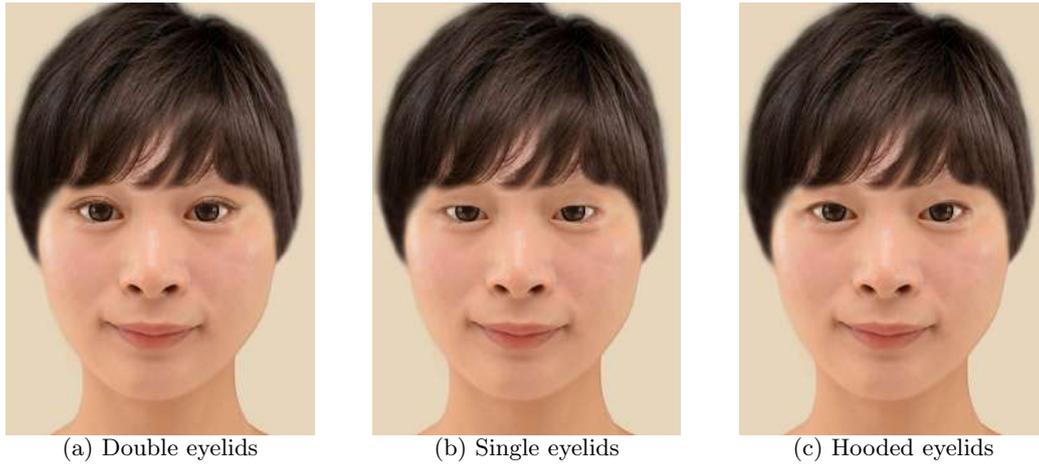


Fig. 5 The face images used in the experiment with different eyelids shapes

Table 1 Eyeliner for length manipulation

Length Eyelid Shapes	1.4mm (5px)	2.8mm (10px)	4.2mm (15px)	5.5mm (20px)	6.9mm (25px)	8.3mm (30px)	9.7mm (35px)	11.1mm (40px)
Single Eyelids								
Hooded Eyelids								
Double Eyelids								

Table 2 Eyeliner for thickness manipulation

Thickness Eyelid Shapes	0.8mm (3px)	1.4mm (5px)	1.9mm (7px)	2.8mm (10px)	3.3mm (12px)	4.2mm (15px)	4.7mm (17px)	5.5mm (20px)
Single Eyelids								
Hooded Eyelids								
Double Eyelids								

adjusting only length or only thickness for a total of six requirements. The following three conditions were given: 1. make the eyeliner that makes the eyes appear the largest, 2. make the eyeliner at the limit of the sense of incongruity, and 3. make the face the most attractive with eyeliner.

We did not set time limits in the experiment, and we suggested that the participants take breaks at their own discretion so that they could make correct decisions. We also instructed them to answer the questions as intuitively as possible.

The trial was conducted in the order of the single eyelids model, the hooded eyelids model, and the double eyelids model. We conducted these trials for three models and provided a short break (at least one-minute) before changing models and conducting the next trial.

After conducting the verification experiment, we surveyed the participants about the criteria for the sense of incongruity and attractiveness. The survey asked for text open-ended questions such as “In which cases did you feel the sense of incongruity?” and “In which cases did you find most attractive?”

In this experiment, we fixed the thickness for adjusting only the length of the eyeliner as the standard thickness and the length for adjusting only the thickness as the standard length. The standard thickness was fixed at 2.8 mm for the single eyelids model, 1.4 mm for the hooded eyelids model and the double eyelids model. For the hooded eyelids model and the double eyelids model, the standard thickness was determined based on 1.5 mm brush tip of a typical ultrafine pencil eyeliner. For the single eyelids model, the standard thickness was determined based on the results of a preliminary experiment conducted by three participants with makeup experience who were not part of the experiment because the shape of the eyeliner is quite different from other models. The standard length was fixed at 2.8 mm for three models. For the double eyelids model, we determined the standard length so that the horizontal coordinate was equal to the point at the end of the double eyelid line on the outer corner of the eye. We adopted the same standard length for the single eyelids model and the hooded eyelids model.

3. Results and Analysis

3.1 Results

The experimental results for the length and the thickness of eyeliner are shown in **Fig. 6** to **Fig. 8** by each eyelids shape using the box plot. The vertical axis indicates the length or thickness of the eyeliner, and the horizontal axis indicates the condition. The larger the value of eyeliner length or thickness, the heavier the eye makeup, and the smaller the value, the lighter the eye makeup. Cross marks in the graph indicate mean values. Then, the mean values for each condition are shown in **Table 3** by each eyelids shape. Moreover, **Fig. 9** to **Fig. 11** show the results of the experiment separately for the experienced makeup and the inexperienced makeup participants. In **Table 3** and **Fig. 6** to **Fig. 11**, “Largest” indicates the eyeliner that makes the eyes appear the largest, “Incongruity” indicates the eyeliner at the limit of the sense of incongruity, and “Attractive” indicates the eyeliner with the most attractive. In **Table 3**, “Largest & Incongruity” shows the smaller value between “Largest” and “Incongruity.”

3.2 Analysis

The purpose of this experiment is to verify whether the eyeliner style that makes the eyes appear the largest without any sense of incongruity is an attractive eyeliner

style. The effect of the eye size illusion becomes larger when the thickness of the eyeliner is thicker. There have been found that larger eyes make attractive, but excessively thick eyeliner is not appropriate as eye makeup because it has any sense of incongruity¹¹⁾. Therefore, we conducted an equivalence test to confirm whether the eyeliner style that makes the eyes appear the largest without any sense of incongruity and the eyeliner style with the most attractive are equivalent. An equivalence test demonstrates statistical equivalence. Equivalence can be proven if the 95% confidence interval of the difference between levels is within a predefined equivalence margin (a range in which no significant difference is determined). In this experiment, the equivalence margin was set to 1.5 mm for both length and thickness because the brush tip of a typical ultra-fine pencil eyeliner is 1.5 mm¹⁷⁾. **Figure 12** shows the results of the equivalence test with the equivalence margin and 95% confidence intervals for the difference between the means of “Largest & Incongruity” and “Attractive” for each eyelids shape. The 95% confidence intervals for all eyelids shapes for both eyeliner length and thickness were within the equivalence margin. The results of the equivalence test showed that there was no significant difference between “the eyeliner style that makes the eyes appear the largest without any sense of incongruity” and “the eyeliner style with the most attractive” for each eyelids shape and indicated that they were equivalent. As a combination of the standard length or the standard thickness, the length of 6.3 mm and the thickness of 2.6 mm for single eyelids, the length of 6.6 mm and the thickness of 3.4 mm for hooded eyelids, and the length of 7.0 mm and the thickness of 2.6 mm for double eyelids were shown to make the eyes appear the largest without any sense of incongruity and also to make the face more attractive (**Table 4**).

The equivalence test was also conducted on both the experienced and the inexperienced participants. The results showed that “the eyeliner style that makes the eyes appear the largest without any sense of incongruity” and “the eyeliner style with the most attractive” except for the length of double eyelids on the experienced participants were equivalent. For the length of double eyelids on the experienced participants, 95% confidence interval was out of the equivalence margin (mean difference 0.74, 95%CI [-0.11, 1.59]).

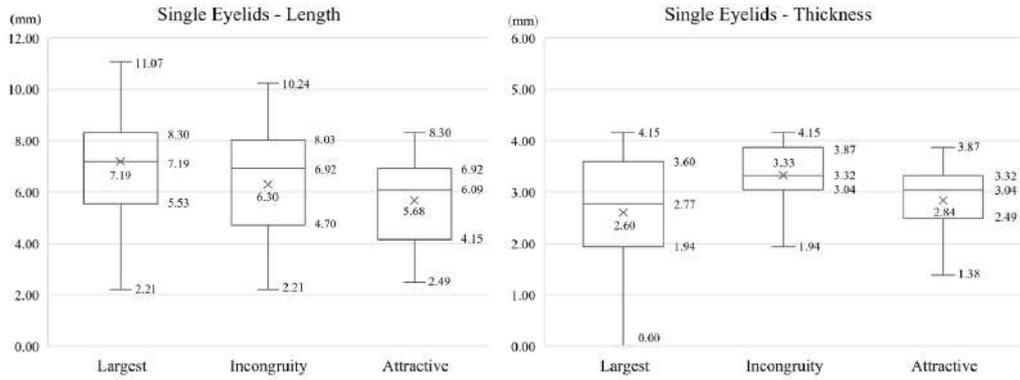


Fig. 6 A result of the single eyelids eyeliner

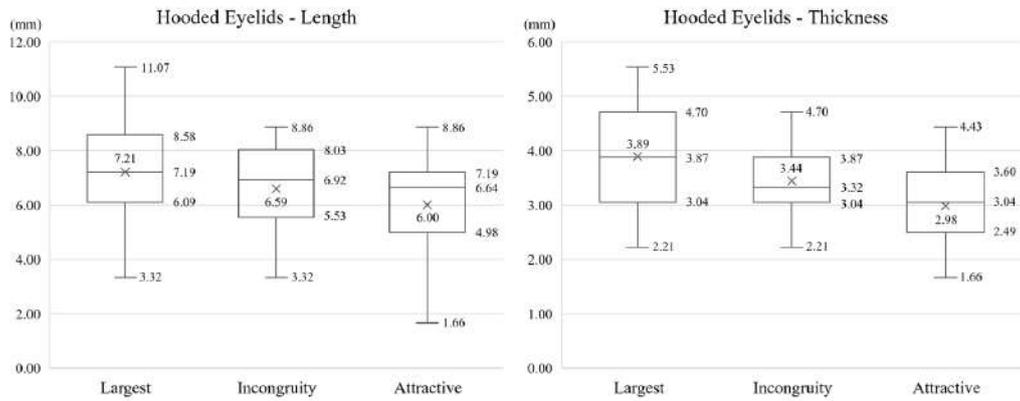


Fig. 7 A result of the hooded eyelids eyeliner

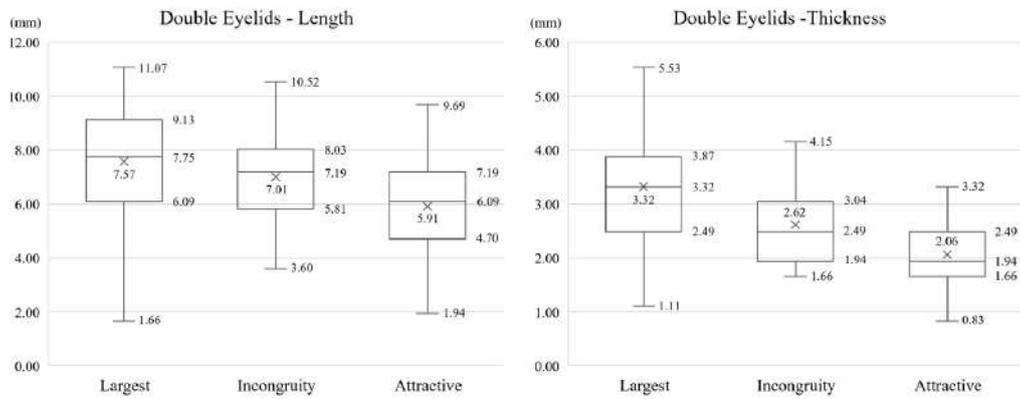


Fig. 8 A result of the double eyelids eyeliner

Table 3 A result of the experiment (mean value)

Eyelid Shapes	Measurements	Conditions			
		Largest	Incongruity	Largest & Incongruity	Attractive
Single Eyelids	Length	7.19	6.30	6.30	5.68
	Thickness	2.60	3.33	2.60	2.84
Hooded Eyelids	Length	7.21	6.59	6.59	6.00
	Thickness	3.89	3.44	3.44	2.98
Double Eyelids	Length	7.57	7.01	7.01	5.91
	Thickness	3.32	2.62	2.62	2.06

(mm)

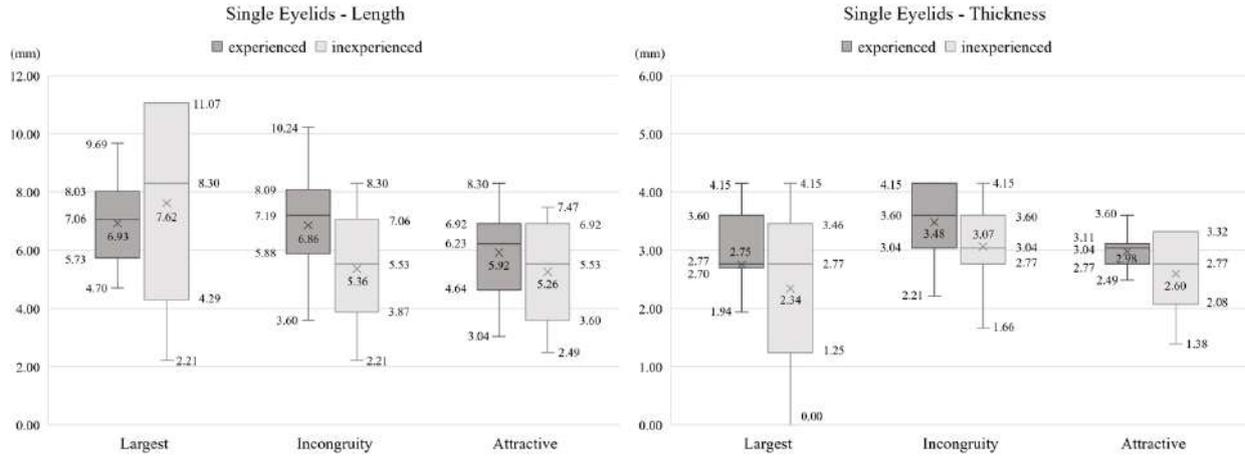


Fig. 9 A result of the single eyelids eyeliner (Comparison of experienced makeup and inexperienced makeup participants)

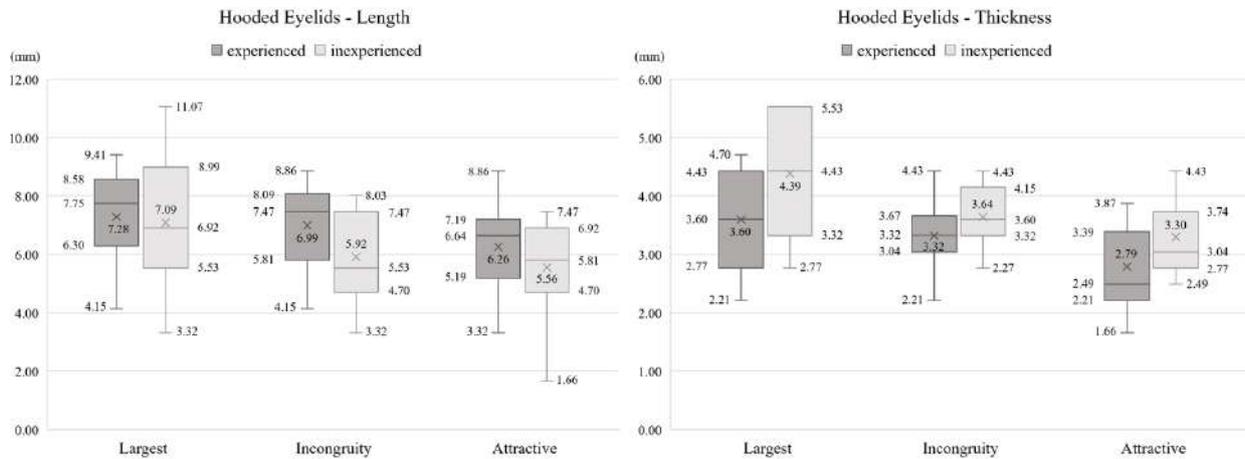


Fig. 10 A result of the hooded eyelids eyeliner (Comparison of experienced makeup and inexperienced makeup participants)

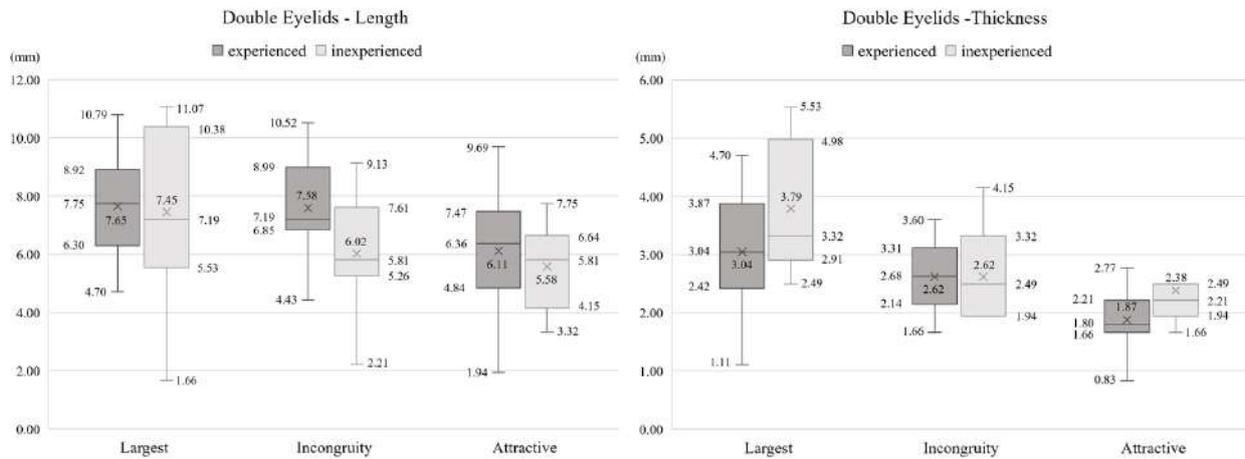


Fig. 11 A result of the double eyelids eyeliner (Comparison of experienced makeup and inexperienced makeup participants)

4. Discussion

4.1 Attractive eyeliner

The equivalence test of the experimental results showed that the eyeliner which makes the eyes appear the largest

without any sense of incongruity is equivalent to the most attractive eyeliner. These results suggest that the eyeliner style that can make the eyes appear the largest within not excessive as makeup is attractive. The post experiment

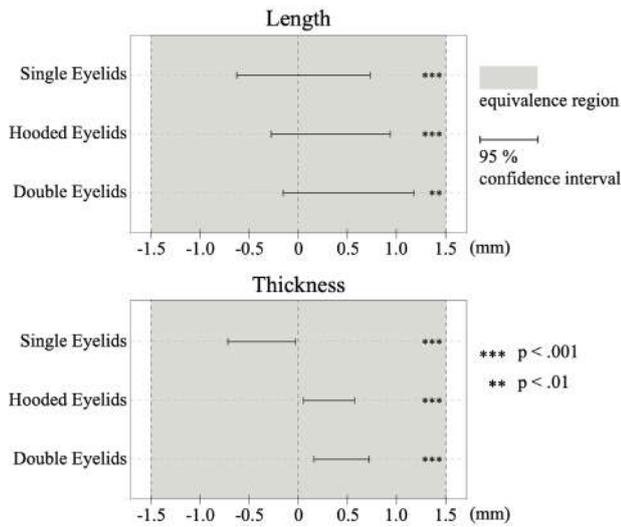


Fig. 12 A result of the equivalence test

survey on attractiveness criteria supported the results of this verification experiment since more than half of the respondents felt that the faces with moderately large eyes were attractive. However, the length of eyeliner which makes the eyes appear the largest without any sense of incongruity on the double eyelids was not equivalent to the most attractive length of eyeliner for the experienced participants. In the case of the double eyelids, the length of the attractive eyeliner seemed to reflect preference rather than making the eyes appear larger.

The eyes with excessive eyeliner make the boundary between the face and the eyes clear, and the geometric illusion caused by the contrast between the surrounding the face and the eyes conversely makes the eyes appear smaller¹⁰⁾. According to the experiment of measuring the eye size illusion by the thickness of eyeliner¹¹⁾, the eyes appeared to be larger up to a certain level of thickness of the eyeliner, but the illusion effect decreased beyond that level. In our experiment, the peak of attractiveness and the peak of appearing to have larger eyes without any sense of incongruity were equivalent. These results suggest that excessive eyeliner not only decreases the illusion of larger eyes but also may not be attractive.

4.2 Differences for eyelids shape

Since the eyeliner model differs significantly between the single and the hooded/double eyelids, we discuss the results separately for the single eyelids and the hooded/double eyelids.

4.2.1 Single eyelids

Figure 9 shows that the IQR are larger for the inexperienced participants than for the experienced participants, indicating that there are larger scatter in the data. The

Table 4 Eyeliner that made the eyes appear the largest without any sense of incongruity

	Length	Thickness
Single Eyelids	 6.3mm	 2.6mm
Hooded Eyelids	 6.6mm	 3.4mm
Double Eyelids	 7.0mm	 2.6mm

experienced participants are familiar with eyeliner as a makeup method, and it is easy for them to set a standard based on the eyeliner they usually wear. However, the inexperienced participants are not familiar with eyeliner, and it is difficult for them to set a standard. In particular, for the single eyelids eyeliner, the eyeliner style was more unfamiliar than other eyeliner styles because it was drawn only on the outer corners of the eyes. This is a reason for the large scatter in the data for the inexperienced participants.

4.2.2 Hooded eyelids and double eyelids

Comparing the results for hooded eyelids with double eyelids (Fig. 7 and Fig. 8), there was a significant difference in “Attractiveness” for the thickness of eyeliner between the double eyelids and the hooded eyelids (Wilcoxon signed rank test, $p < .01$, 95%CI[0.69, 1.11]), with smaller values for the double eyelids. This suggests that the light eyeliner was preferred for double eyelids. In terms of an illusion, the double eyelids alone have an illusion effect of making the eyes appear larger by around 4.5%^{10),14)}. Furthermore, the combination of eyeliner and the double eyelids produces an illusion effect of nearly 6%¹⁴⁾. Since the effect of the illusion on the face⁹⁾⁻¹¹⁾ is not so great compared to a geometric illusion of simple circle combinations, it can be assumed that the effect of the makeup illusion on the face is limited to the extent that it does not become unnatural, and no more the effect can be obtained. The double eyelids already have clear double eyelids line in the vertical direction, so even if the thickness of eyeliner is more modest than the hooded eyelids, the illusion effect could be enough to make the face look more attractive. The post experiment survey showed

that more than half of the respondents felt the sense of incongruity when the eyeliner was thicker than the double eyelids line for hooded and double eyelids. Moreover, some respondents felt that even light eyeliner (eyeliner of modest length and thickness) that clearly shows the double eyelids line for double eyelids was attractive. Therefore, the double eyelids line is an important factor to determine the eyeliner style.

The thickness of eyeliner that makes the eyes appear the largest without any sense of incongruity for double eyelids in this experiment was 2.6 mm. According to the psychophysical measurement¹¹⁾, it was about 2.4 mm. Since the difference of 0.2 mm is less than 1 px, the results of this experiment is approximated as the psychophysical measurement result. However, while Okuri et al.¹¹⁾ reported differences in the illusion effect between male and female participants depending on the thickness of eyeliner, there was no difference between male (the inexperienced participants) and female (the experienced participants) in the “Largest” for the thickness (Fig. 11) of eyeliner in our experiment (Wilcoxon rank sum test, $p = .14$, 95%CI[-1.66, 0.27]). In their study, the illusion effect increased with thicker eyeliner only for male, while the illusion effect did not vary with thicker eyeliner for female. In contrast, female also perceived that the thick eyeliner made the eyes appear larger in our experiment, and the eyeliner that makes the eyes appear the largest was approximately the same for both male and female. We assume that this difference in results is due to the difference in experimental methods. Their experiment used the staircase method to discretely present different level of the length and thickness of the eyeliner. In our approach for using the slider, the user interactively and continuously changes the stimulus so the approach can measure more detailed makeup differences. Moreover, in their experiment, the image presentation time was as short as 3 seconds, whereas in our experiment, the presentation time was not limited, and the participants adjusted the slider while looking at the image carefully. These differences in experimental methods may have some affected the results.

5. Conclusions

In this paper, we conducted a verification experiment in which we quantitatively investigated the eyeliner style that can make the eyes appear the largest without any sense of incongruity and the most attractive eyeliner style for each eyelids shape by using the eyeliner model. The

results of the experiment indicated that the eyeliner that makes the eyes appear the largest without any sense of incongruity is equivalent to the most attractive eyeliner. Furthermore, excessive eyeliner was found to decrease attractiveness. These findings are a contribution to the study of digital makeup.

The eyeliner models were determined from general eyeliner shapes. Therefore, the result obtained in this experiment is an attractive eyeliner in the case of the general shape model. Further experiments are needed in order to obtain findings on other eyeliner shapes. The combination of the length and the thickness in cases where no standard length or thickness has been determined is a future issue. Consideration of eye characteristics such as eye aspect ratio and iris ratio is a future topic.

Eye makeup includes not only eyeliner but also eyeshadow and mascara. Generally, if people wear heavy makeup with thick eyeliner, they also wear other makeup as heavy. Therefore, it is important to consider the total balance of the whole eye makeup. It is a future task to develop methods that can adjust other eye makeup, such as eye shadow and mascara, with parameters and to investigate the total balance of the eye makeup.

In the future, we will consider developing a system that can recommend eye makeup styles that increase attractiveness according to the individual’s eyelids shape based on the results obtained in this paper. Moreover, obtaining feedback from makeup experts would be future tasks.

References

- 1) R. L. Dipboye, R. D. Arvey, D. E. Terpstra: “Sex and Physical Attractiveness of Raters and Applicants as Determinants of Résumé Evaluations”, *Journal of Applied Psychology*, Vol.62, No.3, pp.288-294 (1977).
- 2) J. Y. Baudouin, G. Tiberghien: “Symmetry, Averageness, and Feature Size in the Facial Attractiveness of Women”, *Acta Psychologica*, Vol.117, No.3, pp.313-332 (2004).
- 3) D. Jones, C. L. Brace, W. Jankowiak, K. N. Laland, L. E. Musselman, J. H. Langlois, L. A. Roggman, D. Pérusse, B. Schweder, D. Symons: “Sexual Selection, Physical Attractiveness and Facial Neoteny: Cross-Cultural Evidence and Implications.”, *Current Anthropology*, Vol.36, No.5, pp.723-748 (1995).
- 4) D. I. Perrett, K. J. Lee, Penton-Voak, D. Rowland, S. Yoshikawa, D. M. Burt, S. P. Henzi, D. L. Castles, S. Akamatsu: “Effects of Sexual Dimorphism on Facial Attractiveness”, *Nature*, Vol.394, No.6696, pp.884-887 (1998).
- 5) K. Tagai, H. Ohtaka, H. Nittono: “Faces with Light Makeup Are Better Recognized than Faces with Heavy Makeup”, *Frontiers in Psychology*, Vol.7, No.226 (2016).
- 6) R. Mulhern, G. Fieldman, T. Hussey, J. Lévêque, P. Pineau: “Do Cosmetics Enhance Female Caucasian Facial Attractiveness?”, *International Journal of Cosmetic Science*, Vol.25, No.4, pp.199-205 (2003).

- 7) A. C. Killian, S. Mitra, J. J. Peissig: "The Role of Regional Contrast Changes and Asymmetry in Facial Attractiveness Related to Cosmetic Use", *Frontieres in Psychology*, Vol.9, No.2448, pp.1-11 (2018).
- 8) A. L. Jones, A. Porcheron, R. Russell: "Makeup Changes the Apparent Size of Facial Features", *Psychology of Aesthetics, Creativity, and the Arts*, Vol.12, No.3, pp.359-368 (2018).
- 9) K. Morikawa, S. Matsushita, A. Tomita, H. Yamanami: "A Real-Life Illusion of Assimilation in the Human Face: Eye Size Illusion Caused by Eyebrows and Eye Shadow", *Frontiers in Human Neuroscience*, Vol.9, No.139, pp.1-9 (2015).
- 10) S. Matsushita, K. Morikawa, H. Yamanami: "Measurement of Eye Size Illusion Caused by Eyeliner, Mascara, and Eye Shadow", *Journal of Cosmetic Science*, Vol.66, No.3, pp.161-174 (2015).
- 11) R. Okuri, S. Kodama, T. Takahashi: "Measurement of Eye Size Illusion Caused by Thickness of Eyeliner on Double-Eyelid Eyes", *IIEEJ Transactions on Image Electronics and Visual Computing*, Vol.10, No.1, pp.47-54 (2022).
- 12) H. B. Jung, D. G. Han, J. S. Shim, Y. J. Lee, S. Kim: "Comparison of Eye Measurements between Young Korean Women with Inborn Double Eyelids and Those with Single Eyelids", *Archives of Aesthetic Plastic Surgery*, Vol.26, No.1, pp.7-11 (2020).
- 13) T. Takehara, N. Inoue, L. Yamamoto, M. Shimizu: "Faces with Double Upper Eyelids are Attractive while Lower Eyelids with Eye Bag Makeup might be Unattractive", *Journal of Japan Society of Kansei Engineering*, Vol.20, No.2, pp.121-128 (2021) (in Japanese).
- 14) K. Nishino, N. Okiyama: "A Psychological Study for the Visual Effects Induced by Double Eyelids - Compared with Eyeliner and Considered for Gender Difference", *FRAGRANCE JOURNAL*, Vol.49, No.5, pp.46-52 (2021) (in Japanese).
- 15) C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C. Chang, M. G. Young, J. Lee, W. Chang, W. Hua, M. George, M. Grundmann: "MediaPipe: A Framework for Building Perception Pipelines", *arXiv* (2019).
- 16) M. Kouchi, M. Mochimaru: "Anthropometric Database of Japanese Head 2001", *National Institute of Advanced Industrial Science and Technology*, H16PRO-212 (2008).
- 17) CANMAKE TOKYO, Creamy Touch Liner, <https://www.canmake.com/en/item/detail/36> (2022).

(Received September 9, 2022)

(Revised October 29, 2022)



Rena Okuri

She received her B.E. and M.E degree from Tokyo Denki University in 2020 and 2022. Her research interests include facial image processing.



Shuhei Kodama

He received his Ph.D. (Engineering) from Tokyo Denki University in 2019. He was an engineer at ASTRODESIGN, Inc. from 2019 to 2021. He is currently an assistant professor at Tokyo Denki University. His research fields are NPR, HCI, and fabrication.



Tokiichiro Takahashi (Fellow)

He received his B.E. degrees from Niigata University in 1977, and Ph. D. from the University of Tokyo in 2005, respectively. He had been working at Nippon Telegraph and Telephone Corporation and engaged in Pattern Recognition, Computer Graphics, and Learning Science since 1977. Since 2003, he has been a professor at Tokyo Denki University. He has been a visiting researcher at ASTRODESIGN, Inc. since 2017. His research interests include computer graphics, image processing, and visual computing.

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: Osamu Uchida
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
©2022 IIEEJ