Microscale-searching Evolutionary Optimization for Image Matting

Li Kang, Han Huang*, *Senior Member, IEEE,* Yihui Liang*, *Member, IEEE*, Zeyang Liu, and Zhifeng Hao, *Senior Member, IEEE*

***Abstract*—Image matting is a fundamental computer vision task. It is a large-scale optimization problem involving the selection of pixel pairs from pixel-pair sets of an image. Less prior, such as the trimaps of images, is required for evolutionary algorithms (EAs) to solve the image matting problem compared to deep learning-based methods. However, it is challenging for EAs to solve the image matting problem efficiently due to the large size of the decision set. This paper proposes a framework to guide EAs to search in a microscale subset of the decision set. The subset is estimated by collecting best-so-far solutions during EAs solve similar subproblems. Experimental results demonstrated that by integrating the proposed framework, EAs require fewer FEs to achieve competitive results compared to their original versions. Additionally, the results also indicate that the proposed strategy enhance the performance of EAs in terms of mean squared error and connectivity metrics compared to other EAs-based methods in most cases. The contribution of our work is to make EAs efficient algorithms to solving the image matting problem in scenarios with weak prior.**

***Index Terms*—Image Matting, microscale-searching, large-scale optimization.**

I. Introduction

IMAGE matting is a fundamental task for computer vision, playing a critical role in film editing [1], preprocessing for remote sensing [2] and video postprocessing [3]. It involves generating an alpha matte for an image, where the alpha values, ranging from 0 to 1, represent the opacity of foreground objects. This task can be modeled as a large-scale combinatorial optimization problem, which is about selecting pixel pairs for user-specified pixels. Several types of evolutionary algorithms (EAs), including differential evolution [4], particle swarm evolution [5], [6], multi-objective evolution [7], and competitive swarm optimization [8] have been utilized to solve the image matting problem.

Deep learning-based methods, which utilize deep models like deep neural networks, have recently become the dominant approach to solving image matting [9], [10], while EAs have demonstrated the potential to generate high-quality alpha mattes [4], [7], [11]. Prior is crucial for both EAs-based methods and deep learning-based methods. Most deep learning-based methods use a large amount of finely annotated alpha mattes and trimaps. A finely annotated alpha matte includes labeling foreground pixels with an alpha value of 1, background pixels with an alpha value of 0, and pixels with alpha values between 0 and 1. In contrast, EAs-based methods primarily rely on trimaps as prior. In trimap, users only need to coarsely label unknown pixels for which alpha values need to be predicted, as well as known foreground and background pixels as shown in Fig. 1. Compared to finely annotated alpha mattes, the prior in trimaps is weaker. Finely annotated alpha mattes are vital for enhancing the generalization performance of deep models. In the most popular public dataset [12], the number of pixels with alpha values between 0 and 1 is at least on the order of 10^3. However, it is both time-consuming and labor-intensive to obtain such detailed annotations. Therefore, it is necessary to design efficient EAs-based methods to predict accurate alpha mattes when only weak prior like trimaps is available.

(a) Image (b) Annotated Alpha Matte (c) Trimap

Fig. 1. Comparison between an annotated alpha matte and a trimap. In both trimap and alpha matte, black pixels denote the background, and white pixels indicate the foreground. In the trimap, gray pixels are user-specified and represent represent pixels where the alpha values are to be predicted. In the alpha matte, pixels with alpha values ranging from 0 to 1 are distributed near the foreground boundaries

Optimization-based methods can be categorized into propagation-based [13]–[15], sampling-based [16]–[18], and EAs-based methods [6], [7], [11]. All these approaches require only weak prior to predict alpha mattes. Propagation-

This work is supported by National Natural Science Foundation of China (62276103), Innovation Team Project of General Colleges and Universities in Guangdong Province (2023KCXTD002), the Research and Development Project on Key Technologies for Intelligent Sensing and Analysis of Urban Events Based on Low-Altitude Drones (2024BQ010011), the 2023 Special Program for Audit Theory Research, Guangdong Provincial Philosophy and Social Science Planning (GD23SJZ09), the National Natural Science Foundation of China(62476163,U24A20233), Guangdong Basic and Applied Basic Research Foundation (2023B1515120020), National Key R&D Program of China(2025YFC3410000).

Li Kang, and Zeyang Liu are with the School of Software Engineering, South China University of Technology, Guangzhou 510006, China, Key Laboratory of Big Data and Intelligent Robot (SCUT), MOE of China, Guangzhou 510006, China. (e-mail: k_conley@foxmail.com, 202020145275@scut.edu.cn).

Han Huang is with Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China and Guangzhou Zhisuan Linghang Technology Co., Ltd.(hanhuang@foxmail.com)

Y. Liang is with the Zhongshan Institute, University of Electronic Science and Technology of China, Zhongshan 528400, China (e-mail: yihuiliang@outlook com).

Zhifeng Hao is with the Department of Mathematics, College of Science, Shantou University, Shantou 515063, China (e-mail: haozhifeng@stu.edu.cn).

based methods model the image matting problem as a linear system, considering relations between local [13] or non-local pixels [14], [15]. These methods aim to solve the problem in closed-form to obtain accurate alpha mattes, assuming a linear relationship between pixel colors or features. However, the accuracy of obtained alpha mattes may decrease when this linear relationship is not met, which is common in natural images.

Different from the propagation-based methods, the sampling-based methods focus on finding optimal pixel pairs to compute accurate alpha mattes for unknown pixels. These methods typically select a subset of pixel pairs that are expected to contain optimal pixel pairs for one or multiple unknown pixels. The subset selection process can be modeled as a single objective problem, such as a sparse subset selection problem [17]. A small number of pixel sets of foreground and background are selected by measuring the KL-divergence between the feature vectors of the unknown and known pixel sets. The subset selection process can be also modeled as a multi-objective optimization problem [7], [18] to alleviate the influence of the conflict among different sampling criteria. Brute-force methods are employed to determine the best pixel pair in the subset [17], [18] because the size of the subsets is small. Additionally, a region-based random walk algorithm is proposed to solve the image matting problem when all known foreground and background pixels on the boundary of the unknown region are included in the subset [16]. Furthermore, the subset of pixel pairs can serve as the basis for establishing surrogate models for the evaluation of pixel pairs [19]. However, the main limitation of sampling-based methods is the rapid decline in the the accuracy of alpha mattes when the pixel pairs in the subset have large difference from optimal pixel pairs [18].

In EAs-based methods, the image matting problem is modeled as a large-scale optimization problem about selecting optimal pixel pairs for unknown pixels. The dimensionality of the problem is twice the number of unknown pixels which are decided by trimaps. As indicated in the CEC'2010 competition on large-scale global optimization [20], the dimensionality of a large-scale optimization problem is more than 10^3. For instance, the number of unknown pixels of a 480×800 image in the benchmark dataset [12] is approximately 10^4. Therefore, the image matting problem is a large-scale optimization problem. Cooperation co-evolutionary framework is used to solve the problem [4], [7], [8] as it has been proven to be effective in solving solve large-scale optimization problems [21]–[23]. The problem is decomposed into multiple subproblems under the framework. The existing EAs-based methods concentrate on the allocation of fitness evaluations (FEs) among subproblems to improve the efficiency of EAs. The allocation is mainly guided by the similarity among subproblems [4], [8], [11]. The convergence speed of EAs is also measured to control the allocation of FEs [5]. Nevertheless, the enhancement in the efficiency of EAs through the allocation of FEs is constrained because some solutions found for subproblems have limited contribution to the search for the optimal solution of the whole problem.

During the process of optimizing the image matting problem, some solutions are evaluated repetitively across different subproblems. Let $\mathcal{V}$ denote the set that comprises solutions repetitively evaluated for multiple subproblems. Solutions in $\mathcal{V}$ are evaluated repetitively by EAs for different subproblems, but their contribution to the optimization of the whole problem are probably the same or close due to the prevalence of solutions with the same fitness value for similar subproblems [7]. The accumulation of repetitive FEs escalates if the number of similar subproblems are large. The increase in these repetitive FEs diminishes the probability of finding potential high quality solutions due to the finite number of available FEs. Consequently, it is crucial to reduce the number of the repetitively evaluated solutions in $\mathcal{V}$ to improve the efficiency of EAs.

This paper proposes a framework for EAs to solve the image matting problem where a microscale subset of the decision set is estimated for EAs to find the best solutions. The microscale subset of the decision set is estimated by incorporating the optimization of similar subproblems. EAs explore the subset to find the best solutions with a scheme of dynamic allocation of FEs. The contribution of our work is to make EAs an efficient approach to solving the image matting problem in weak prior scenarios.

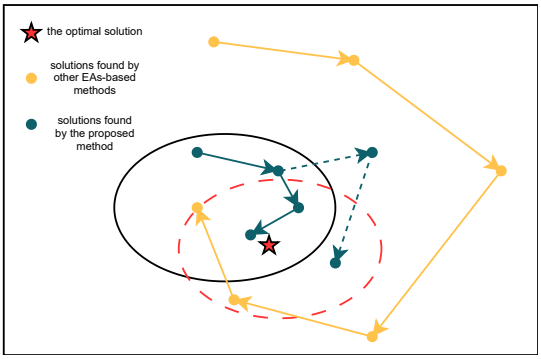


Fig. 2. Comparison of the process of searching for solutions between EAs with the proposed method and their original versions. The rectangle is a decision set. The red and black ellipses denote the subset of the definition and the subset found by our strategy, respectively

The remainder of this paper is organized as follows. The mathematical model of image matting is demonstrated in Section II. The microscale-searching optimization algorithm is introduced in Section III. A framework based on microscale-searching for EAs is proposed in Section IV. Empirical results are presented and discussed in Section V. Section VI concludes this paper.

II. Large-scale Combinatorial Optimization Model of Image Matting

This section introduces the large-scale combinatorial optimization model of image matting. From the perspective of image matting [24], [25], the color vector I_z of unknown pixel z can be modeled as a linear combination of a pixel pair, which can be modeled as

$$I_z = \alpha_z F_z + (1 - \alpha_z) B_z$$

where $\alpha_z \in [0,1]$ is the alpha value of z. F_z and B_z are vectors of RGB color of a selected pixel pair. As shown in Fig. 3, the unknown pixels, the set of foreground and background pixels are given by a trimap. Given a selected pixel pair, the alpha value can be calculated by

$$\hat{\alpha}_z = \frac{(I_z - B_z)(F_z - B_z)}{||F_z - B_z||^2} \tag{1}$$

where $\hat{\alpha}_z$ is the estimated alpha value for z.

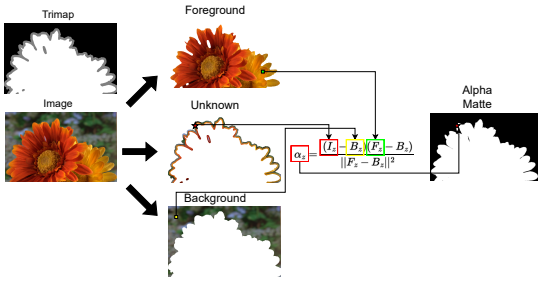


Fig. 3. Procedure of calculating alpha value for a pixel. The resolution of image is 497×800. The image involves 2×10^5 foreground pixels, 1×10^5 background pixels, and 3×10^4 unknown pixels.

If a pixel pair is selected for z, the difference between the true alpha value α_z^* and the estimated alpha value $\hat{\alpha}_z$ computed by Eq. (1) is approximated by a handcrafted fitness function g_z because α_z^* is unknown. Let Ω_F, Ω_B, and Ω_U denote the sets of foreground, background pixels, respectively. The optimal pixel pairs selection problem can be modeled as

$$\min\ G(X) = \sum_{i=1}^{N} g_{z_i}(x_{F,z_i}, x_{B,z_i}) \tag{2}$$

$$s.t.\ X = \begin{pmatrix} x_{F,z_1}, & x_{F,z_2}, & ..., & x_{F,z_N}, \\ x_{B,z_1}, & x_{B,z_2}, & ..., & x_{B,z_N} \end{pmatrix} \tag{3}$$

$$x_{F,z_i} \in \Omega_F, x_{B,z_i} \in \Omega_B \tag{4}$$

$$z_i \in \Omega_U, i = 1, 2, ..., N \tag{5}$$

$$\Omega_F \cup \Omega_B \cup \Omega_U = T \tag{6}$$

where the decision variables x_{F,z_i} and x_{B,z_i} denote the indexes of selected foreground and background pixels for the unknown pixel z_i. N denotes the total number of unknown pixels in Ω_U. The range of values for x_{F,z_i} and x_{B,z_i} are restricted to the set of positive integers, which characterizes this problem as a combinatorial optimization problem.

Both the number of foreground pixels and background pixels of the image in Fig. 3 are about 10^5 according to the corresponding trimap. For one unknown pixel, the number of pixel pairs that can be selected is equal to the size of the combination of foreground pixel set and background pixel set. The size of the combination is about 10^{10}. According to the CEC'2010 competition on large-scale global optimization [20], image matting is a large-scale optimization problem because the number of unknown pixels is more than 10^3. The image matting problem is still challenging due to the following reasons: 1) the handcrafted function g_{z_i} is non-convex and multi-peak [7]. 2) The dimensionality of the decision vector is high and the value ranges of the decision variables are large. 3) The image matting problem is not fully separable because x_{F,z_i} interacts with x_{B,z_i}.

EAs have been applied to solve various computer vision tasks efficiently [26]–[28]. Although some optimization methods like brute-force [17], [29] and random walk [16] have been used to solve image matting, it is difficult for them to utilize similarity of subproblems to improve the solving efficiency due to the large size of decision sets. The population-based characteristic provides advantages for EAs to solve large-scale optimization problems [21], [30], [31]. However, the efficiency of EAs in solving image matting is influenced by repetitive FEs.

III. Microscale Valid Decision Subset for Large-scale Optimization Problem of Image Matting

This section presents a strategy to estimate a microscale subset of the decision set where the number of repetitively evaluated solutions is reduced for EAs. We assume that FEs can be reduced from the perspective of subsets of the decision set. The definition of valid decision subsets is introduced based on the assumption. A strategy is proposed to estimate valid decision subsets from disjoint subsets of pixel pairs based on the definition. Finally, an analysis is presented to show that the valid decision subset found by the proposed strategy is a microscale subset of the decision set.

A. Assumption of Microscale Searching

This subsection presents an assumption about the reduction of FEs from the perspective of subsets of the decision set. Let $\Omega = \Omega_F \times \Omega_B$ denote a set of pixel pairs. The Cartesian product of N sets Ω is the decision set D of the optimal pixel pairs selection problem.

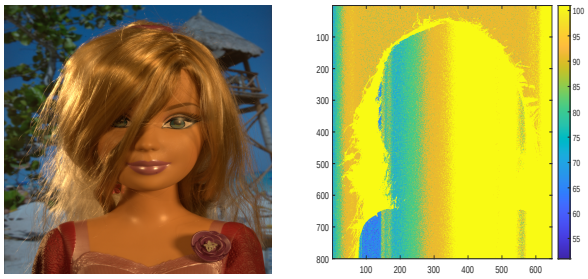
Definition 1 (decision subset): A decision subset V is a subset of decision set D that consists of feasible solutions found by algorithms.

The optimal pixel pairs selection problem can be decomposed into $1 \leq K \leq N$ subproblems since the objective function (Eq. (2)) is the sum of N function values. If $K = N$, the decision set of each subproblem is Ω. Let $V_k \subseteq \Omega$ and $V_z \subseteq \Omega$ denote decision subsets of the subproblem k and the subproblem z respectively. If the pixel k and the pixel z are similar, the intersection of V_k and V_z could cause repetitive FEs because solutions that have the same value of fitness functions for similar pixels widely exist in the decision set [7]. As shown in Fig. 4, many pixel pairs are evaluated multiple times for a given number of FEs.

Let A denote the decision vector of a subproblem. The decision set D_A of A is the Cartesian product of N_s sets Ω. $N_s \leq N$ is the number of similar pixels. If $N_s = 1$, $D_A = \Omega$. If $N_s = N$, $D_A = D$.

Definition 2 (valid decision subset): Let $\mathcal{V}$ denote a decision subset of D_A. $\mathcal{V}$ is a valid decision subset if $\mathcal{V}$ satisfies the following conditions,

$$d(A^*, \hat{A}) < d(A^*, \overline{A}) \quad \forall \hat{A} \in \mathcal{V}, \overline{A} \notin \mathcal{V} \tag{7}$$



(a) (b)

Fig. 4. Visualization of the number of repetitive FEs. Brighter pixels are evaluated more often in (b)

where A^* is the optimal solution of A. $\overline{A}$ is any solution out of $\mathcal{V}$. A function d is used to calculate the difference between the fitness values of two solutions.

According to *Definition* 2, although solutions in $\mathcal{V}$ may lead to repetitive FEs, they contribute more significantly to the overall solution compared to solutions not in $\mathcal{V}$. It means that the efficiency of EAs can be improved if EAs can be guided to search in a microscale valid decision subset.

In the following subsections, an algorithm is proposed to estimate a microscale valid decision subset $\mathcal{V}$. A dynamic allocation strategy of FEs is proposed based on the similarity of subproblems to find the best solutions based on $\mathcal{V}$.

B. Strategy for Estimating a Microscale Valid Decision Subset

This subsection proposes a microscale-searching strategy to estimate a valid decision subset. For the i-th and the k-th subproblems, their fitness functions satisfy

$$\exists \epsilon \geq 0, |g_i(\hat{x}) - g_k(\hat{x})| \leq \epsilon \tag{8}$$

where $\hat{x}$ is a selected pixel pair. The maximum of ϵ reflects how the i-th subproblem is similar to the k-th subproblems. If ϵ is equal to 0, the i-th and k-th subproblems can be seen as the same. Let $\mathcal{V}_i$ denote a valid decision subset of the i-th subproblem. x_i^* is the optimal solution of the i-th subproblem. $\overline{x}$ is any solution out of $\mathcal{V}_i$. According to *Definition* 2, the relations of the fitness values of $\hat{x}$, x_i^*, and $\overline{x}$ satisfies

$$|g_i(\hat{x}) - g_i(x_i^*)| < |g_i(\overline{x}) - g_i(x_i^*)| \quad \forall \hat{x} \in \mathcal{V}_i, \overline{x} \notin \mathcal{V}_i \tag{9}$$

Let $\mathcal{V}_k$ denote a valid decision subset of the k-th subproblem. Eq. (9) can be rewritten as follows by substituting Eq. (8) into Eq. (9)

$$|g_k(\hat{x}) - g_i(x_i^*)| < |g_i(\overline{x}) - g_i(x_i^*)| + \hat{\epsilon} \quad \forall \hat{x} \in \mathcal{V}_k, \overline{x} \notin \mathcal{V}_k \tag{10}$$

$\hat{\epsilon}$ is the maximum of ϵ that makes Eq. (10) true. (10) indicates that $\mathcal{V}_i$ can be estimated by $\mathcal{V}_k$. If the value of $\hat{\epsilon}$ is small, the estimated subset gets close to the true subset because $\hat{\epsilon}$ represents the similarity of the two subproblems. On the other hand, a small size of the estimated subset implies that $\hat{\epsilon}$ is small because $\hat{x}$ needs to be close to the optimal solution. This implies that we can enhance the quality of V_k subject to the constraint of its scale.

A strategy is proposed to solve the microscale valid decision subset estimation problem by incorporating the optimization of similar subproblems (**Algorithm 1**). The estimation of a valid decision subset is decomposed into a search for its elements. During each iteration, the best solution in the population is added into the estimated valid decision subset. The strong global search capacity of evolutionary algorithms ensures that these best solutions are winners against other solutions in various areas of the decision set. Eq. (10) dictates that these solutions belong to one of optimal valid decision subsets. There are $\lceil e/(N_p \cdot t_1) \rceil$ iterations during the optimization (Line 4). r_i controls the generation number of the population in each iteration. The available generation number is $\lfloor t_1 \cdot r_j \rfloor$ (Line 9). Let S_j and H_j denote the sets of best-so-far solutions found in the current iteration and in all previous iterations, respectively. Let $(X_{SF}, Y_{SF}), (X_{SB}, Y_{SB})$ denote the spatial coordinates of pixel pairs in S_j. Let $(X_{HF}, Y_{HF}), (X_{HB}, Y_{HB})$ denote the spatial coordinates of pixel pairs in H_j. r_j is updated as follows

$$r_j = 1 - \frac{1}{2}\left(\frac{|X_{SF} \cap X_{HF}|}{|X_{SF}|} + \frac{|X_{HB} \cap X_{SB}|}{|X_{SB}|}\right) \tag{11}$$

If r_j is equal to 1, all best-so-far solutions found in this iteration are better than those in H_j. FEs are encouraged to allocate to search for better solutions in the next iteration. Otherwise, FEs are preserved to be used to search in the estimated valid decision subset as r_j is close to 0. The valid decision subset obtained by the union of all H_j can be used for all similar subproblems (Line 15).

Algorithm 1 Estimation of Valid Decision Subsets

Input: the number of similar subproblems N_s,the number of FEs e,the population size N_p,t_1
Output: the valid decision subset,
rest FEs for each subproblem

1: Initialize population $P_i, i = 1, 2, ..., N_s$ randomly
2: Initialize $r_i, i = 1, 2, ..., N_s$ as 1
3: Initialize $a_i, i = 1, 2, ..., N_s$ as 0
4: **for** $i = 1$ to $\lceil e/(N_p \cdot t_1) \rceil$ **do**
5: **for** $j = 1$ to N_s **do**
6: $S_j \leftarrow \emptyset$
7: $X_F, Y_F, X_B, Y_B \leftarrow \emptyset$
8: $a_j \leftarrow a_j + (1 - r_j) \cdot N_p \cdot t_1$
9: Obtain solutions s with the best fitness values from each generation while EAs update the population using $\lfloor N_p \cdot t_1 \cdot r_j \rfloor$ fitness evaluations.
10: $S_j \leftarrow S_j \cup \{s\}$
11: Update r_j by Eq. (11)
12: $H_j \leftarrow H_j \cup S_j$
13: **end for**
14: **end for**
15: $\mathcal{V} \leftarrow \bigcup_{i=1}^{N_s} H_i$
16: **return** $\mathcal{V}, a_1, a_2, ..., a_{N_s}$

C. Analysis of the Estimated Microscale Valid Decision Subset

This subsection provides an analysis of the size of valid decision subsets estimated by **Algorithm 1**. Let N_p denote the population size. e/N_p is the maximum number of best-so-far solutions that EAs can find for a subproblem. The actual number of best-so-far solutions is equal to the size of H_i. The size of H_i is smaller than e/N_p because the allocation of FEs is dynamical. The size of H_i satisfies

$$|H_i| \leq e/N_p \tag{12}$$

The size of the valid decision subset for a group of N_s similar subproblems is

$$|\mathcal{V}| = \sum_{i=1}^{N_s} |H_i| \leq (N_s \cdot e/N_p)^{N_s} \tag{13}$$

The size of the decision set of the optimal pixel pair optimization problem is $|\Omega|^{N_s}$ according to Eq. (2). Let ϵ_s denote the ratio of the size of valid decision subsets to that of the decision set. ϵ_s satisfies

$$\epsilon_s \leq (\frac{N_s \cdot e/N_p}{|\Omega|})^{N_s} \tag{14}$$

In practice, the number of iterations e/N_p and similar subproblems are far less than $|\Omega|$. The value of ϵ_s will become much smaller than 1 as the number of pixels N_s increases. Therefore, the valid decision subset $\mathcal{V}$ estimated by **Algorithm 1** is microscale. According to *Definition* (2), solutions in the valid decision subset are close to the optimal solution. The rest FEs can be allocated to searching for the best solutions in $\mathcal{V}$.

IV. Framework of Microscale-searching Evolutionary Optimization for Image Matting

This section presents a framework for EAs to solve the image matting problem based on the concept of microscale-searching. The flowchart of the framework is shown in Fig. 5. Unknown pixels are first grouped into M clusters by the agglomerative clustering method [32] based on their features. These features are represented by a contatenation of the pixel's spatial coordinate vector and its RGB color vector. Microscale valid decision subsets are estimated for each group by **Algorithm 1**. **Algorithm 2** is proposed to search for solutions in the estimated microscale valid decision subset. Finally, an alpha matte is calculated by Eq. (1) with the found pixel pairs.

A scheme for dynamically assigning FEs to search for the best solutions in estimated microscale valid decision subsets is proposed. It is indicated that the allocation of FEs should be dynamically adjusted according the contribution of stagnant subpopulations [33]–[35]. The main reason why subpopulations are stagnant is that solutions in $\mathcal{V}$ do not satisfy Eq. (7). Not all subproblems in the same group have similar solutions because it is difficult to extract robust features to measure the correlation between the similarity of subproblems and the similarity of solutions. Let SP denote a set of subproblems that have solutions similar to those of other subproblems. Let $\overline{SP}$ denote a set of subproblems that have no such solutions. A portion of FEs should be allocated to searching for solutions out of $\mathcal{V}$ for $\overline{SP}$ because $\mathcal{V}$ may not involve solutions that satisfy Eq. (7) for $\overline{SP}$. For SP, the majority of FEs should be allocated to searching in $\mathcal{V}$.

The correlation between the similarity of subproblems and the similarity of solutions is calculated by the distance from the

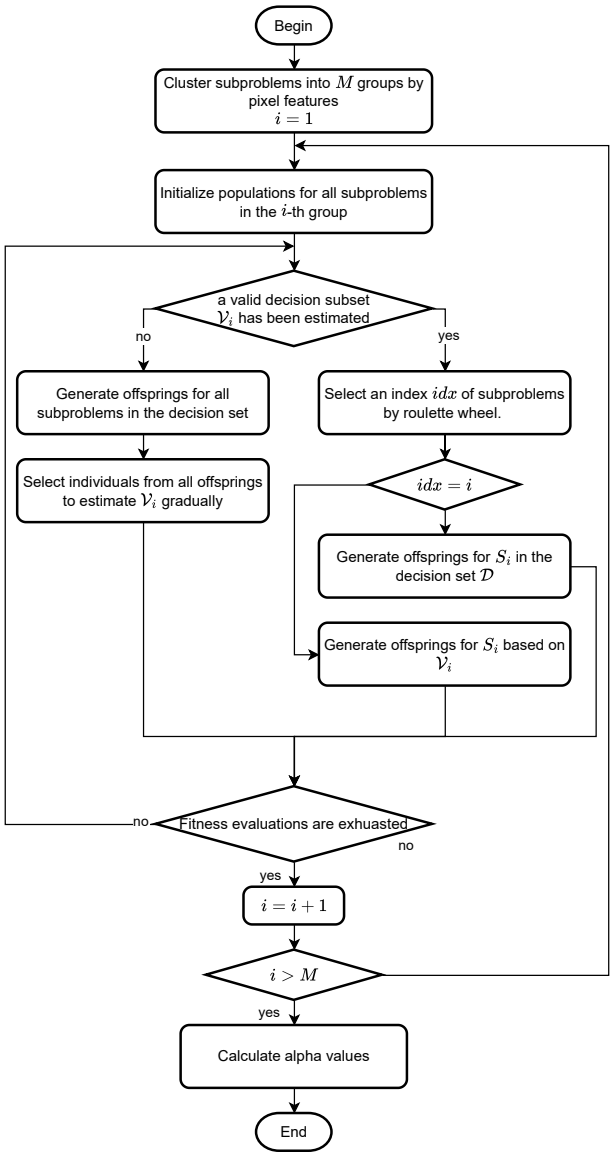


Fig. 5. Flowchart of microscale-searching evolutionary optimization for image matting.

feature vectors of subproblems to the centroid of the feature vectors. The correlation can be calculated as follows

$$\overline{f} = \frac{1}{N_s}\sum_{i=1}^{N_s} f_i, w_i = \frac{||f_i - \overline{f}||_2}{\sum_{j=1}^{N_s} ||f_j - \overline{f}||_2} \tag{15}$$

$$W = (w_1, w_2, ..., w_{N_s}) \tag{16}$$

where f_i is the feature vector of i-th subproblem, and the centroid $\overline{f}$ is calculated by averaging the feature vectors. N_s is the number of similar subproblems. An N_s-dimensional vector W is used to show the correlation among subproblems. w_i is the value of the i-th dimension of the vector. w_i is close to 1 for the subproblems that have feature vectors far from the centroid.

A roulette-wheel selection method is used to decide whether FEs are allocated to searching in $\mathcal{V}$ or not based on W (Line 8). More FEs are allocated to searching out of $\mathcal{V}$ for $\overline{SP}$ than SP because $\overline{SP}$ are more likely to be chosen (Lines

Algorithm 2 Scheme of Allocation of FEs among Different Subproblems

Input: the numbers of FEs $e_1, e_2, ..., e_{N_s}$,the weight vector W, the estimated valid decision subset $\mathcal{V}$, the population size N_p,the decision set D,t_2
Output: the indexes of pixel pairs U

1: Initialize populations $P_i, Q_i, i = 1, 2, ..., N_s$ by randomly selecting solutions in $\mathcal{V}$
2: Initialize $x_i, i = 1, 2, ..., N_s$ as two-dimensional zero vectors
3: **while** $e_i > 0, i = 1, 2, ..., N_s$ **do**
4: **for** $i = 1$ to N_s **do**
5: **if** $e_i \leq 0$ **then**
6: continue
7: **end if**
8: $idx \leftarrow$ the index of a subproblem selected by using roulette-wheel selection based on W
9: **if** $i == idx$ **then**
10: Update P_i by searching in D for t_2 generations
11: **else**
12: Update Q_i by searching in $\mathcal{V}$ for t_2 generations
13: **end if**
14: Obtain the best solution u_{best} by comparing solutions in P_i and Q_i
15: **if** u_{best} has better fitness value than that of x_i **then**
16: $x_i \leftarrow u_{best}$
17: **end if**
18: $e_i \leftarrow e_i - t_2 \cdot N_p$
19: **end for**
20: **end while**
21: **return** $x_1, x_2, ..., x_{N_s}$

9-13). Best solution are obtained by comparing solutions of two populations (Line 14).

The microscale-searching evolutionary optimization as illustrated in Fig. 5 comprises two algorithms. The time complexity of **Algorithm 1** in a single iteration is $O(N_s \times N_p \times G_1)$, where N_s is the number of subproblems, N_p is the population size, and G_1 is the maximum number of generations for **Algorithm 1**. The time complexity of **Algorithm 2** is $O(N_s \times N_p \times G_2)$, where G_2 is the maximum number of generations for **Algorithm 2**. The two algorithms are executed sequentially. Their time complexities can therefore be summed. The complexity for each iteration of the loop is thus $O(N_s \times N_p \times (G_1 + G_2))$. Consequently, the overall time complexity of microscale-searching evolutionary optimization method is $O(M \times N_s \times N_p \times (G_1 + G_2))$.

V. Experiments

This section comprises three experiments. The first experiment shows the comparison of different methods on convergence trends or matting metrics. The second experiment visualizes repetitively evaluated solutions, demonstrating the impact of searching within a microscale valid decision subset on reducing the number of repetitively evaluated solutions. The third experiment investigates the existence of a microscale valid decision subset within the decision set of the image matting problem.

A. Experimental Setup

To evaluate the performance of EAs with MSEO, all experiments are carried out on a benchmark dataset [12] which contains 35 natural images. 27 of all images are training images with ground-truth alpha mattes, while the remaining eight images whose ground-truth is unavailable are provided for evaluation and ranking only.

We use genetic algorithm (GA) [36], competitive swarm optimization (CSO) [30], and particle swarm optimization (PSO) [37] in experiments due to their popularity in literature. MSEO with GA, CSO, and PSO are dubbed MSEO-GA, MSEO-CSO, and MSEO-PSO, respectively. The objective function for individuals in CSO-based approaches is formulated as follows

$$g(x_F, x_B, k) = h_1(x_F, x_B, k) + h_2(x_F, k) + h_3(x_B, k)$$

The objective function evaluates the quality of a pixel pair (x_F, x_B) for the k-th unknown pixel. $h_1(x_F, x_B, k)$ is color chromatic criteria, $h_2(x_F, k)$ is spatial closeness criteria for foreground pixels, and $h_3(x_B, k)$ is spatial closeness criteria for background pixels. The formulations of three criteria are presented as

$$h_1(x_F, x_B, k) = \|C_k - \hat{\alpha} C_{x_F} - (1 - \hat{\alpha}) C_{x_B}\|^2$$
$$h_2(x_F, k) = \|S_k - S_{x_F}\|^2$$
$$h_3(x_B, k) = \|S_k - S_{x_B}\|^2.$$

C_k and S_k are the color vector and the spatial vector of the k-th unknown pixel.

FEs for each subproblem E is set to 5000. Population size N_p is set to 50. The paramters t_1, t_2 in **Algorithm 1** and **Algorithm 2** are set to $20, 10$, $30, 10$, and $30, 10$ for MSEO-GA, MSEO-CSO, MSEO-PSO ,respectively, according to parameters study in Section V.F.

B. Improvement of Efficiency by Microscale-searching on Image Matting

This subsection demonstrates the advantages of MSEO from the perspectives of convergence trends, fitness values, and matting metrics. Mean square error (MSE) as an image matting metric is used to study the accuracy of image matting. MSE measures the difference between the estimated alpha value and the true value [12]. All experiments are run 30 times. Wilcoxon rank sum test with 0.05 significance level is conducted on the experimental results.

The comparison of convergence trends are shown in Fig. 6 to investigate how MSEO cooperates with EAs. PSO, CSO, and GA are involved in this comparison. The vertical axes of curves are log values of fitness values, and the horizontal axes are the number of FEs executed by EAs. The dash lines in the figures are the results of the original EAs. The dot lines are the results of EAs with MSEO. The diamonds on the dot lines indicate that the fitness values are equal to the fitness values of the best solutions found by the original EAs.

There are several sudden drops on the curves. The drops indicate that EAs start to search for the best solutions in the estimated valid decision subsets for different subproblems due to the FEs consumption strategy in **Algorithm 1**. The red diamonds in the convergence trends of Fig. 6 are used to mark the time when MSEO-EAs finds the solutions that have the closest fitness value to that of the best solutions found by original

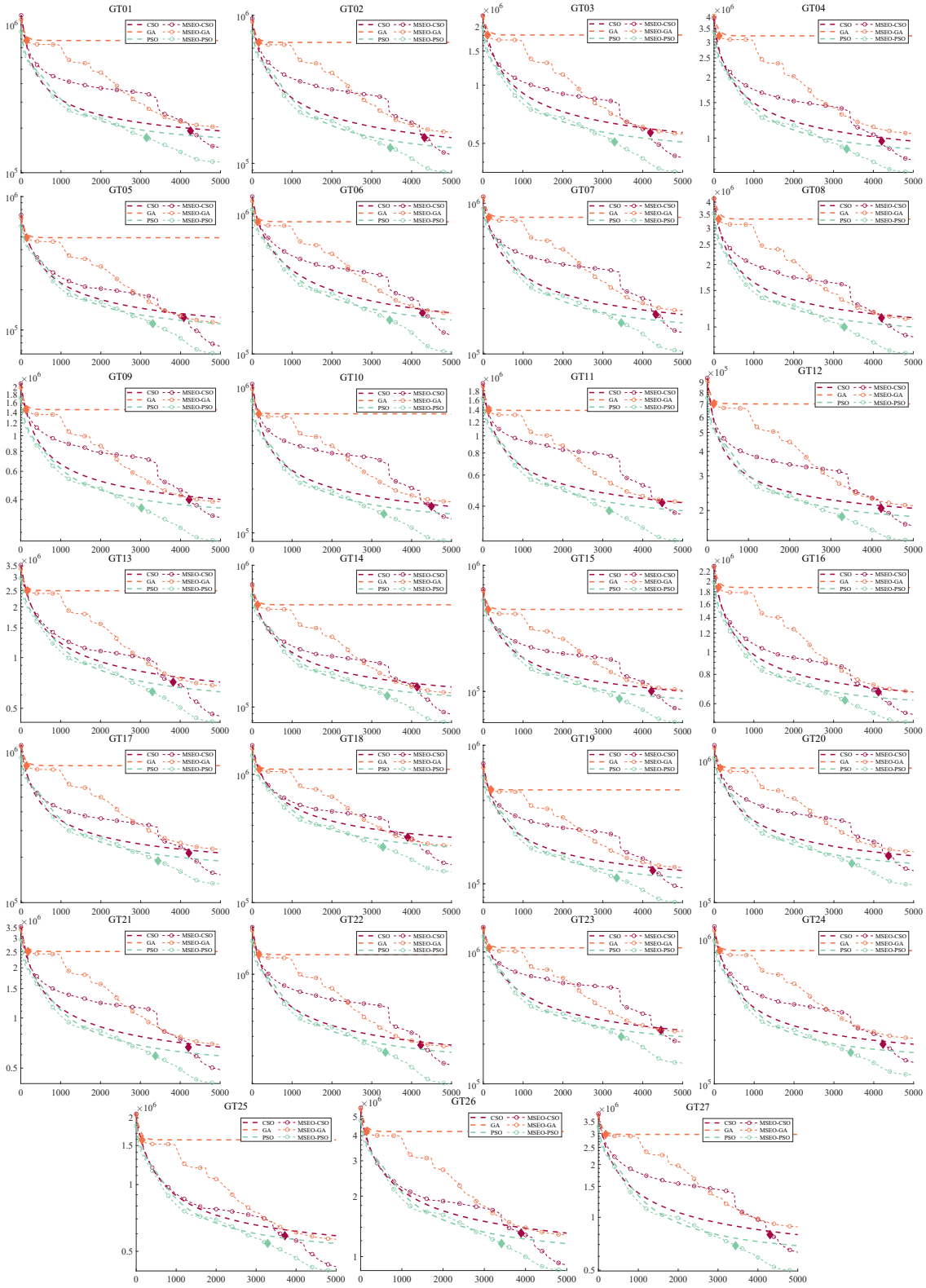


Fig. 6. The vertical axes of curves are fitness value, and the horizontal axes are the number of FEs executed by EAs. The dash lines in the figures are the results of the original EAs. The dot lines are the results of EAs with MSEO. The diamonds on the dot lines indicate that the fitness values are equal to the fitness values of the best solutions found by the original EAs.

EAs. In all examples, GA converges the earliest among all methods. Crossover and mutation operators in GA are directly applicable to the discrete decision set of the optimal pixel pair selection problem. However, these operators neglect the spatial smoothness characteristic of pixels within an image. This characteristic could otherwise provide heuristic information to the algorithm. The incorporation of MSEO supplements GA with this smoothness characteristic. This is achieved because the valid decision subset, integral to MSEO, is estimated based on subproblem similarity. Solutions to similar subproblems tend to be spatially proximate within the image during the optimization process. Conversely, the decision vector for the optimal pixel pair selection problem is first converted to a continuous representation when methods like PSO and CSO optimize this problem. The characteristics among pixels in the image can be more easily utilized in the continuous real-number space to facilitate the search process of EAs. However, for methods like PSO and CSO, although they can find better solutions more easily compared to GA, solutions would be repetitively evaluated across similar subproblems. The exploration efficiency of these methods is compromised by these repetitive FEs. The incorporation of MSEO reduces the exploration cost for these methods in large-scale decision sets, as they can search within the valid decision subset. The valid decision subset has already saved heuristic information from the optimization of similar subproblems across diverse regions of the decision set.

Table I summarizes the comparison of the fitness values of the best solutions and matting metrics of the predicted alpha mattes obtained by MSEO-EAs and original EAs. "$+$" and "$\downarrow$" respectively denote that the results of EAs with MSEO are significantly better to those of the original EAs according to the Wilcoxon test, conducted at a 5% significance level over 30 independent trials. "$-$" signifies results that are significantly worse, while "$\approx$" indicates no significant difference from the compared method. It can be seen that MSEO-EAs find better solutions than the original EAs in all cases. However, the results of MSEO-EAs on matting metrics do not surpass those of the original EAs in some cases, because the fitness function is an approximation of matting metrics. Table II presents the comparison results between EAs combined with MSEO and existing single-objective EAs-based matting methods in terms of four matting metrics, MSE, SAD, CON, and GRAD. MSE and SAD are used to measure the difference between the found pixel pairs and the optimal pixel pairs. GRAD and CON evaluate the visual perceptual differences between the predicted alpha mattes and the ground truth alpha mattes. The comparisons include MSEO-CSO versus GC-CSO [8] and MSEO-PSO versus ACSC-PSO [6]. GC-CSO and ACSC-PSO focus on accelerating the solving of the matting problem using EAs. GC-CSO employs solutions of representative subproblems as solutions for other similar subproblems. ACSC-PSO controls the diversity of solutions by monitoring convergence speed. Our method not only enhances the solution efficiency of evolutionary algorithms but also discovers better solutions because it is based on an effective decision subset.

C. Performance on Weak-prior Scenarios

This subsection demonstrates the advantages of our method by comparing the performance of matting approaches requiring weak priors with that of approaches requiring strong priors on weak-prior scenarios. A weak-prior scenario is defined as one where the data contains trimaps but lacks finely annotated alpha mattes. Existing matting methods that require strong priors are mostly trained using natural image matting datasets. We construct the weak-prior scenario by combining data from natural images and medical images. The selected medical matting dataset is Brain-growth [38] that consists of 39 low-intensity contrast T2-W MR images for the newborn brain's white matter tissue myelination process, and the natural image dataset is AlphaMatting. Since the natural image dataset contains fewer images than the medical image dataset, the weak-prior dataset is composed of 27 randomly selected images from Brain-growth combined with all images from the AlphaMatting dataset.

The methods requiring weak priors included in this experiment are as follows: 1) Propagation-based methods: Closed-Form [13], KNN [15], and Information-Flow [39]. 2) Sampling-based methods: Bayesian Matting [40] and PDMS [18]. The methods requiring strong priors include DIM [41], FBA [42], DiffMatte [43], and MedicalMatting [44]. Specifically, DIM, FBA, and DiffMatte were trained on Composition-1K dataset [41], while MedicalMatting was trained on the Brain-growth dataset.

Table III presents the mean and standard deviation of matting metrics obtained from 10 independent runs. The MSE results in the table demonstrate that the proposed method achieves the most accurate matting masks in weak-prior scenarios. The SAD results are not as strong as the MSE results because the MSE metric tends to amplify predictions that deviate significantly from the ground truth values. In weak-prior scenarios, despite a distribution shift between medical and natural images, both types of images contain similar local regions. This similarity enables deep learning-based methods to achieve highly accurate predictions in these areas. However, in regions dissimilar to natural images, deep learning-based methods struggle to produce accurate predictions, and the MSE metric subsequently amplifies these inaccuracies. The results in the table also reveal that matting methods requiring weak priors generally yield more accurate alpha mattes predictions compared to methods requiring strong priors. This observation underscores the necessity of designing efficient matting methods based on EAs.

D. Effect on the Reduction of Repetitive FEs

Fig. 7 visualizes the number of repetitive FEs for GT01, GT11, and GT27. The first and the second columns are natural images and corresponding trimaps. The visualization results of repetitive FEs are shown in the third, fourth, and fifth rows for CSO, GC-CSO, and MSEO-CSO repetitively. Bright pixels are evaluated more often. The upper bound of the number of repetitive FEs is set differently for the sake of distinction.

As can be seen in the last three columns of Fig. 7, the results of MSEO-CSO have smaller numbers of bright pixels than

TABLE I

"↓" INDICATES THAT THE FITNESS VALUE OF THE BEST SOLUTION OBTAINED BY MSEO-EAS IS SIGNIFICANTLY LESS THAN THAT OF THE BEST SOLUTION OBTAINED BY THE ORIGINAL EAS, WHILE "↑" SIGNIFIES THE OPPOSITE. "+", "−", AND "≈" INDICATE THAT IMPROVING APPROACHES ARE SIGNIFICANTLY BETTER THAN THE ORIGINAL EAS. THE NUMBERS IN THE LAST COLUMN REPRESENT THE TOTAL NUMBER OF "↓" AND "+", RESPECTIVELY.

	GT01	GT02	GT03	GT04	GT05	GT06	GT07	GT08	GT09	GT10	GT11	GT12	GT13	GT14	
CSO	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / -	↓ / +	↓ / -	↓ / -	↓ / +	↓ / ≈	↓ / +	↓ / -	
GA	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	
PSO	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / -	↓ / +	↓ / -	↓ / -	↓ / +	↓ / +	↓ / +	↓ / -	
	GT15	GT16	GT17	GT18	GT19	GT20	GT21	GT22	GT23	GT24	GT25	GT26	GT27		
CSO	↓ / +	↓ / -	↓ / -	↓ / +	↓ / +	↓ / +	↓ / +	↓ / ≈	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	26/19	
GA	↓ / +	↓ / ≈	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	↓ / +	26/26	
PSO	↓ / +	↓ / -	↓ / -	↓ / +	↓ / +	↓ / ≈	↓ / +	↓ / +	↓ / ≈	↓ / +	↓ / +	↓ / +	↓ / +	26/19	

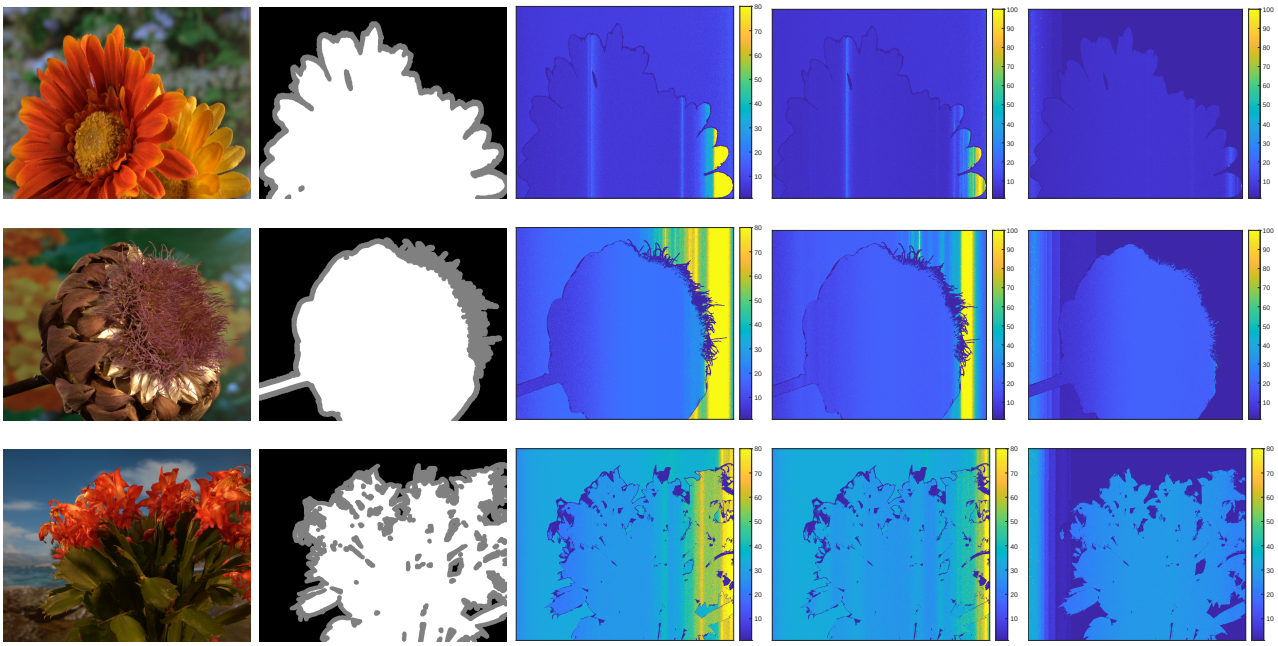


Fig. 7. Visualization results of the number of repetitive FEs of CSO-based methods for GT01, GT07, and GT21 from the first row to the third row. The last three columns visualize the number of repetitive FEs. Bright pixels are evaluated more often.

TABLE II

COMPARISON WITH EXISTING SINGLE-OBJECTIVE EAS-BASED MATTING METHODS IN TERMS OF FOUR MATTING METRICS. THE NUMBERS UNDER EACH METRIC FROM LEFT TO RIGHT REPRESENT THE TOTAL COUNTS WHERE THE EAS COMBINED WITH MSEO IS SIGNIFICANTLY BETTER, SIMILAR, AND SIGNIFICANTLY WORSE THAN OTHER METHODS, RESPECTIVELY.

	MSE	SAD	CON	GRAD
GC-CSO	18/3/6	23/1/3	21/0/6	21/2/4
ASCS-PSO	27/0/0	27/0/0	27/0/0	27/0/0

those of other methods because a lot of repetitively evaluated solutions are avoided during the estimation of microscale valid decision subsets. For MSEO-CSO, most repetitive FEs are generated when EAs search for solutions in the microscale valid decision subset. For CSO, the number of bright pixels is larger than that of other methods because a lot of solutions are repetitively evaluated. For GC-CSO, the number of bright pixels is smaller than that of CSO because the search space is limited by the initial solutions that are the solutions of representative subproblems. The Fig. 6 illustrates that the quality of solutions found by MSEO-CSO initially lags behind that of solutions found by CSO. However, solutions found by MSEO-CSO progressively improve to surpass the quality of those of CSO before the exhaustion of FEs. This progression suggests that the reduction of FEs enable CSO to discover high-quality solutions using fewer FEs. Furthermore, the results in Table I demonstrate that MSEO-CSO also predicts more accurate alpha mattes in all three cases.

E. Existence of Microscale Valid Decision Subsets

The properties of decision subsets found by **Algorithm 1** are presented in this subsection to verify whether a microscale valid decision subset exists in the decision set of the image matting problem. According to *Definition* (2), the difference among fitness values of solutions in a microscale valid decision subset and the optimal solution is smaller than that among solutions out of a microscale valid decision subset and the optimal solution. The proportion of optimal pixel pairs is used to manifest the quality of solutions in the decision subsets. The

TABLE III
COMPARISON OF MATTING METRICS BETWEEN METHODS REQUIRING STRONG PRIORS AND THOSE REQUIRING WEAK PRIORS IN WEAK-PRIOR SCENARIOS. THE TABLE PRESENTS THE MEAN AND STANDARD DEVIATION OF MATTING METRICS FROM 10 INDEPENDENT EXPERIMENTS, WITH THE BEST VALUES HIGHLIGHTED IN BOLD.

Methods	MSE	SAD	CON	GRAD
DIM	0.0530±0.0012	2.2397±0.0073	2.4806±0.0136	2.1518±0.0076
FBA	0.0725±0.0031	1.7979±0.0120	1.7552±0.0253	1.7624±0.0122
MedicalMatting	0.3746±0.0036	22.0245±0.0151	20.8340±0.0356	22.0070±0.0153
DiffMatte	0.0870±0.0050	**1.6750**±0.0155	**1.6449**±0.0157	**1.5941**±0.0247
Closed-Form	0.0500±0.0013	2.9354±0.0058	5.0067±0.0134	2.8837±0.0062
KNN	0.0564±**0.0007**	3.6230±**0.0050**	4.1823±**0.0098**	3.5249±0.0054
Information-Flow	0.0464±0.0013	2.7059±0.0054	2.5809±0.0103	2.6365±0.0057
Bayesian	0.1548±0.0016	14.8008±0.0065	17.5435±0.0114	14.9725±0.0067
PDMS	0.0443±0.0016	2.7769±0.0057	2.6824±0.0131	2.6692±0.0060
MSEO-CSO	**0.0431**±0.0013	2.5544±0.0063	2.0628±0.0107	2.0664±**0.0000**

TABLE IV
PROPERTIES OF VALID DECISION SUBSETS. "DS" DENOTE DECISION SETS. "DSS" DENOTE DECISION SUBSETS. BOTH THE SIZES OF "DS" AND "DSS" ARE EQUAL TO THE NUMBER OF PIXEL PAIRS. $|\cdot|$ MEANS THE SIZE OF THE SET.

Image IDs	**GT01**	**GT02**	**GT03**	**GT04**	**GT05**	**GT06**	**GT07**	**GT08**	**GT09**
$\lvert DS\rvert$	3.02E+154	1.76E+150	9.00E+220	1.77E+337	2.06E+303	4.08E+215	3.83E+201	7.97E+663	2.83E+781
$\lvert DSS\rvert$	2.20E+74	1.60E+72	2.10E+108	3.90E+176	9.00E+162	8.50E+107	3.60E+96	8.60E+369	1.30E+440
ϵ_s	7.29E-81	9.10E-80	2.33E-112	2.21E-161	4.38E-141	2.08E-108	9.39E-106	1.08E-294	4.59E-342
$\hat{\epsilon}_s$	6.30E-26	1.80E-26	3.90E-37	4.50E-59	9.10E-52	4.90E-36	3.30E-34	8.80E-115	2.10E-136
τ	91.1%	99.7%	98.9%	99.7%	95.1%	98.9%	99.1%	97.5%	96.3%
Image IDs	**GT10**	**GT11**	**GT12**	**GT13**	**GT14**	**GT15**	**GT16**	**GT17**	**GT18**
$\lvert DS\rvert$	2.33E+318	9.10E+157	2.02E+286	8.56E+243	1.80E+185	3.27E+263	6.98E+588	3.83E+254	8.37E+193
$\lvert DSS\rvert$	6.70E+167	2.30E+72	1.50E+145	1.40E+127	8.50E+92	3.90E+136	3.30E+343	5.10E+128	1.00E+98
ϵ_s	2.88E-152	2.53E-86	7.44E-142	1.64E-117	4.72E-93	1.19E-127	4.73E-247	1.33E-125	1.19E-96
$\hat{\epsilon}_s$	2.80E-55	3.30E-26	1.60E-50	8.70E-41	4.60E-32	1.20E-46	2.30E-103	9.60E-43	9.90E-32
τ	99.4%	99.7%	96.7%	98.8%	99.4%	99.6%	98.3%	99.5%	99.7%
Image IDs	**GT19**	**GT20**	**GT21**	**GT22**	**GT23**	**GT24**	**GT25**	**GT26**	**GT27**
$\lvert DS\rvert$	3.45E+119	5.93E+337	9.35E+252	9.43E+246	6.12E+282	1.42E+533	5.73E+866	1.34E+806	2.73E+660
$\lvert DSS\rvert$	1.40E+55	2.50E+175	5.10E+127	8.40E+121	5.80E+137	5.20E+306	1.00E+530	1.40E+479	1.30E+364
ϵ_s	4.06E-65	4.21E-163	5.46E-126	8.91E-126	9.48E-146	3.66E-226	1.75E-336	1.05E-328	4.76E-296
$\hat{\epsilon}_s$	2.00E-20	2.20E-58	1.10E-43	4.70E-41	1.00E-48	5.60E-93	5.60E-156	4.20E-143	1.60E-115
τ	99.5%	99.1%	99.7%	98.2%	99.7%	98.8%	99.3%	99.2%	99.7%

proportion τ of optimal pixel pairs in the decision subsets is calculated as

$$\tau = \frac{1}{m}\frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i} o_{i,j}}{\sum_{i=1}^{m} n_i}$$

$$o_{i,j} = \begin{cases} 1, & \text{if an optimal pixel pair exists for } j\text{-th} \\ & \text{subproblem in } i\text{-th cluster} \\ 0, & \text{otherwise} \end{cases}$$

n_i is the number of subproblems in i-th cluster. The quality of solutions in the decision subsets is proportional to τ because the difference between the fitness values of best-so-far solutions and the optimal solution becomes smaller if the optimal pixel pairs are contained in the decision subsets. The mean sizes of decision subsets found for all clusters is used to calculate ϵ_s of Eq (14) to show the scale of decision subsets compared to the size of the decision set.

The first row of Table IV shows the mean sizes of the decision set of all clusters. The second row shows the mean sizes of decision subsets of all clusters. The third row shows the ϵ_s that is the result of the division of $|DSS|$ and $|DS|$. The fourth row shows the $\hat{\epsilon}_s$ calculated by Eq. (14) when E is 5000 and N is 50. The size of decision subsets found by **Algorithm 1** is microscale because ϵ_s is smaller than $\hat{\epsilon}_s$ for all cases. The last row of Table IV shows that more than 95 per cent of optimal pixel pairs of subproblems are contained in decision subsets. Although not all optimal pixel pairs are the optimal solutions of the objective function because the matting equation Eq. (1) is ill-posed, the fitness values of optimal pixel pairs are better than that of most pixel pairs. The result of Table IV indicates that microscale valid decision subsets are involved in decision subsets found by **Algorithm 1** for most cases. The reasons that some optimal pixel pairs are not included in valid decision subsets are attributed as follows. 1) The handcrafted image features mislead clustering algorithms to incorrectly group a few subproblems. 2) Although repetitive FEs are avoided, the number of FEs is still not enough for EAs to find optimal solutions.

F. Parameter Analysis

The performance of MSEO is mainly influenced by the subset estimated by **Algorithm 1**. t_1, t_2 are investigated in Table V on GT01 because they are vital for **Algorithm 1**. t_2 is first fixed to study the influence of t_1 because the solutions in estimated microscale valid decision subsets is determined by t_1. t_2 is studied after the best value of t_1 is set. t_1 control the number of generations in one iteration. A small t_1 would

TABLE V
COMPARISON OF MSE, CON, AND FITNESS VALUES AMONG DIFFERENT SETTING OF (t_1, t_2). BOLDED TEXT INDICATES THE BEST RESULTS.

MSEO-CSO					MSEO-GA					MSEO-PSO				
t_1	t_2	MSE	CON	Fitness	t_1	t_2	MSE	CON	Fitness	t_1	t_2	MSE	CON	Fitness
10	10	1.91E-03	4.63E+02	1.46E+05	10	10	2.21E-03	5.35E+02	1.87E+05	10	10	1.93E-03	4.60E+02	1.41E+05
20	10	1.83E-03	4.35E+02	1.22E+05	**20**	**10**	**2.12E-03**	**5.16E+02**	**1.95E+05**	**20**	**10**	**1.79E-03**	**4.23E+02**	**1.15E+05**
30	**10**	**1.78E-03**	**4.19E+02**	**1.22E+05**	30	10	2.17E-03	5.18E+02	2.32E+05	30	10	1.79E-03	4.26E+02	1.17E+05
40	10	1.97E-03	4.58E+02	1.83E+05	40	10	5.35E-03	7.27E+02	5.37E+05	40	10	1.93E-03	4.50E+02	1.55E+05
50	10	5.56E-03	7.41E+02	1.91E+05	50	10	5.56E-03	7.41E+02	7.34E+05	50	10	5.56E-03	7.41E+02	1.78E+05
30	20	1.82E-03	4.34E+02	1.23E+05	20	20	2.14E-03	5.13E+02	2.06E+05	20	20	1.78E-03	4.24E+02	1.12E+05
30	30	1.81E-03	4.30E+02	1.25E+05	20	30	2.17E-03	5.18E+02	2.14E+05	20	30	1.78E-03	4.22E+02	1.13E+05
30	40	1.81E-03	4.33E+02	1.27E+05	20	40	2.15E-03	5.14E+02	2.20E+05	20	40	1.80E-03	4.23E+02	1.15E+05
30	50	1.81E-03	4.25E+02	1.27E+05	20	50	2.23E-03	5.39E+02	2.31E+05	20	50	1.81E-03	4.28E+02	1.15E+05

cause EAs to stop prematurely. A large t_1 would result in few number of FEs left for searching in the estimated valid decision subsets. This is consistent with the results in the third to seventh rows of Table V. t_2 is related to the number of times that EAs search out of the estimated valid decision subsets. The eighth row to the last row of Table V indicates that whether EAs search outside the subsets has less impact on the results after the valid decision subsets are estimated. The results of Table also validates the existence and effectiveness of valid decision subsets.

VI. DISCUSSION

The experimental results validate that our proposed Microscale-searching Evolutionary Optimization (MSEO) framework effectively enhances the efficiency and quality of EA-based image matting. The primary strength of our approach lies in its robust performance in scenarios with weak priors, where deep learning methods, reliant on extensive high-quality training data, often struggle. By operating as a single-instance optimization, our algorithm requires no pre-training. The microscale-searching mechanism successfully balances global exploration with local exploitation, allowing the EA to search high-quality solutions with fewer fitness evaluations compared to its standard counterparts. This positions our method as a powerful tool for applications demanding high-fidelity matting without access to large, labeled datasets or precise user-provided trimaps.

Despite its advantages, we acknowledge the inherent limitations of our framework, which in turn motivate future work. The most significant trade-off is the computational cost. While more efficient than traditional EAs, our optimization-based approach is slower than the inference speed of pre-trained deep learning models, making it better suited for offline processing rather than real-time applications. Furthermore, the framework's core mechanism assumes a degree of structural self-similarity within the image, and its effectiveness may be reduced on highly heterogeneous images. Consequently, future research will focus on two key areas: 1) accelerating the optimization process, potentially through GPU parallelization or hybrid algorithms, to broaden its applicability; and 2) developing more adaptive knowledge transfer mechanisms that are less dependent on spatial proximity and can handle a wider variety of image content. Addressing these limitations will be crucial in advancing EA-based methods as a practical and powerful alternative in the image matting toolkit.

VII. CONCLUSION

In this paper, we propose a microscale-searching framework to improve the efficiency of EAs to solve the image matting problem. A microscale valid decision subset is estimated by collecting best-so-far solutions of similar subproblems. Competitive solutions can be found by EAs with the help of the proposed microscale-searching evolutionary optimization with less number of FEs than other EAs-based image matting methods. The proposed microscale-searching framework make EAs efficient approaches to solving the image matting problem. Experimental results validate that the estimated subset is substantially smaller than the decision set. For more than 90% of the subproblems, the estimated subsets also include the optimal solution. The existence of microscale valid decision subsets provides guidance for a reasonable allocation of FEs. The framework FEs allocation strategy improves the efficiency of EAs in solving the image matting problem.

In the future, we will try to design more efficient EAs-based approaches based on the concept of microscale valid decision subsets for other computer vision tasks.

REFERENCES

[1] M. Gong, Y. Qian, and L. Cheng, "Integrated foreground segmentation and boundary matting for live videos," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1356–1370, 2015.

[2] W. Li, Z. Zou, and Z. Shi, "Deep matting for cloud detection in remote sensing images," *IEEE Trans. Geosci. Remote. Sens.*, vol. 58, no. 12, pp. 8490–8502, 2020.

[3] B. Zhu, Y. Chen, J. Wang, S. Liu, B. Zhang, and M. Tang, "Fast deep matting for portrait animation on mobile phone," in *Proc. ACM Int. Conf. Multimedia*, 2017, pp. 297–305.

[4] Z. Cai, L. Lv, H. Huang, H. Hu, and Y. Liang, "Improving sampling-based image matting with cooperative coevolution differential evolution algorithm," *Soft Comput.*, vol. 21, no. 15, pp. 4417–4430, 2017.

[5] Y. Liang, H. Huang, Z. Cai, and L. Lv, "Particle swarm optimization with convergence speed controller for sampling-based image matting," in *Int. Conf. Intelligent Computing*, July 2018, pp. 656–668.

[6] Y. Liang, H. Huang, and Z. Cai, "PSO-ACSC: a large-scale evolutionary algorithm for image matting," *Frontiers Comput. Sci.*, vol. 14, no. 6, p. 146321, 2020.

[7] Y. Liang, H. Huang, Z. Cai, and Z. Hao, "Multiobjective evolutionary optimization based on fuzzy multicriteria evaluation and decomposition for image matting," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 5, pp. 1100–1111, 2019.

[8] F.-J. Feng, H. Huang, Q.-X. Wu, X. Ling, Y. Liang, and Z. Cai, "An alpha matting algorithm based on collaborative swarm optimization for high-resolution images," *Scientia Sinica Informationis*, vol. 50, no. 3, pp. 424–437, 2020.

[9] J. Li, J. Zhang, S. J. Maybank, and D. Tao, "Bridging composite and real: Towards end-to-end deep image matting," *Int. J. Comput. Vis.*, vol. 130, no. 2, pp. 246–266, 2022.

[10] Y. Qiao, Y. Liu, Z. Wei, Y. Wang, Q. Cai, G. Zhang, and X. Yang, "Hierarchical and progressive image matting," *ACM Trans. Multim. Comput. Commun. Appl.*, vol. 19, no. 2, pp. 52:1–52:23, 2023.

[11] F. Feng, H. Gou, Y. Liang, L. Feng, M. Tan, H. Huang, and L. Wang, "Micro-scale searching algorithm for high-resolution image matting," *Multim. Tools Appl.*, vol. 83, no. 13, pp. 38 931–38 947, 2024.

[12] C. Rhemann, C. Rother, J. Wang, M. Gelautz, P. Kohli, and P. Rott, "A perceptually motivated online benchmark for image matting," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2009, pp. 1826–1833.

[13] A. Levin, D. Lischinski, and Y. Weiss, "A closed-form solution to natural image matting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 228–242, 2008.

[14] P. G. Lee and Y. Wu, "Nonlocal matting," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2011, pp. 2193–2200.

[15] Q. Chen, D. Li, and C. Tang, "KNN matting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2175–2188, 2013.

[16] K. He, C. Rhemann, C. Rother, X. Tang, and J. Sun, "A global sampling method for alpha matting," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2011, pp. 2049–2056.

[17] L. Karacan, A. Erdem, and E. Erdem, "Alpha matting with kl-divergence-based sparse sampling," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4523–4536, 2017.

[18] H. Huang, Y. Liang, X. Yang, and Z. Hao, "Pixel-level discrete multiobjective sampling for image matting," *IEEE Trans. Image Process.*, vol. 28, no. 8, pp. 3739–3751, 2019.

[19] Y. Liang, H. Gou, F. Feng, G. Liu, and H. Huang, "Natural image matting based on surrogate model," *Appl. Soft Comput.*, vol. 143, p. 110407, 2023.

[20] K. Tang, X. Li, P. N. Suganthan, Z. Yang, and T. Weise, "Benchmark functions for the cec'2010 special session and competition on large-scale global optimization," *Nature inspired computation and applications laboratory*, vol. 24, pp. 1–18, 2007.

[21] Z. Yang, K. Tang, and X. Yao, "Large scale evolutionary optimization using cooperative coevolution," *Inf. Sci.*, vol. 178, no. 15, pp. 2985–2999, 2008.

[22] X. Li and X. Yao, "Cooperatively coevolving particle swarms for large scale optimization," *IEEE Trans. Evol. Comput.*, vol. 16, no. 2, pp. 210–224, 2012.

[23] M. N. Omidvar, X. Li, Y. Mei, and X. Yao, "Cooperative co-evolution with differential grouping for large scale optimization," *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 378–393, 2014.

[24] A. R. Smith and J. F. Blinn, "Blue screen matting," in *Proc. Annu. Conf. Comput. Graph. Interactive Tech.*, 1996, pp. 259–268.

[25] T. K. Porter and T. Duff, "Compositing digital images," in *Proc. Annu. Conf. Comput. Graph. Interactive Tech.*, 1984, pp. 253–259.

[26] Y. Bi, B. Xue, and M. Zhang, "Genetic programming with image-related operators and a flexible program structure for feature learning in image classification," *IEEE Trans. Evol. Comput.*, vol. 25, no. 1, pp. 87–101, 2021.

[27] J. Wei, G. Zhu, Z. Fan, J. Liu, Y. Rong, J. Mo, W. Li, and X. Chen, "Genetic u-net: Automatically designed deep networks for retinal vessel segmentation using a genetic algorithm," *IEEE Trans. Medical Imaging*, vol. 41, no. 2, pp. 292–307, 2022.

[28] F. Zhao, Z. Zeng, H. Liu, R. Lan, and J. Fan, "Semisupervised approach to surrogate-assisted multiobjective kernel intuitionistic fuzzy clustering algorithm for color image segmentation," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 6, pp. 1023–1034, 2020.

[29] E. S. Varnousfaderani and D. Rajan, "Weighted color and texture sample selection for image matting," *IEEE Trans. Image Process.*, vol. 22, no. 11, pp. 4260–4270, 2013.

[30] R. Cheng and Y. Jin, "A competitive swarm optimizer for large scale optimization," *IEEE Trans. Cybern.*, vol. 45, no. 2, pp. 191–204, 2015.

[31] M. N. Omidvar, Y. Mei, and X. Li, "Effective decomposition of large-scale separable continuous functions for cooperative co-evolutionary algorithms," in *IEEE Congr. Evol. Computation*, 2014, pp. 1305–1312.

[32] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. Prentice-Hall, 1988.

[33] P. Xu, W. Luo, X. Lin, Y. Chang, and K. Tang, "Difficulty and contribution-based cooperative coevolution for large-scale optimization," *IEEE Trans. Evol. Comput.*, vol. 27, no. 5, pp. 1355–1369, 2023.

[34] M. N. Omidvar, B. Kazimipour, X. Li, and X. Yao, "CBCC3 - A contribution-based cooperative co-evolutionary algorithm with improved exploration/exploitation balance," in *IEEE Congr. Evol. Computation*, 2016, pp. 3541–3548.

[35] M. Yang, A. Zhou, C. Li, J. Guan, and X. Yan, "CCFR2: A more efficient cooperative co-evolutionary framework for large-scale global optimization," *Inf. Sci.*, vol. 512, pp. 64–79, 2020.

[36] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press, 1992.

[37] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.* IEEE, 1995, pp. 1942–1948.

[38] H. B. Li, F. Navarro, and e. a. Ivan Ezhov, "QUBIQ: uncertainty quantification for biomedical image segmentation challenge," *CoRR*, vol. abs/2405.18435, 2024.

[39] Y. Aksoy, T. O. Aydin, and M. Pollefeys, "Designing effective inter-pixel information flow for natural image matting," in *CVPR*. IEEE Computer Society, 2017, pp. 228–236.

[40] Y. Chuang, B. Curless, D. Salesin, and R. Szeliski, "A bayesian approach to digital matting," in *CVPR (2)*. IEEE Computer Society, 2001, pp. 264–271.

[41] N. Xu, B. L. Price, S. Cohen, and T. S. Huang, "Deep image matting," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2017, pp. 311–320.

[42] M. Forte and F. Pitié, "F, b, alpha matting," *CoRR*, vol. abs/2003.07711, 2020.

[43] Y. Hu, Y. Lin, W. Wang, Y. Zhao, Y. Wei, and H. Shi, "Diffusion for natural image matting," in *ECCV (57)*, ser. Lecture Notes in Computer Science, vol. 15115. Springer, 2024, pp. 181–199.

[44] L. Wang, X. Ye, L. Ju, W. He, D. Zhang, X. Wang, Y. Huang, W. Feng, K. Song, and Z. Ge, "Medical matting: Medical image segmentation with uncertainty from the matting perspective," *Comput. Biol. Medicine*, vol. 158, p. 106714, 2023.

Li Kang received the B.S. degree in electronic science and technology from Shenzhen University, China, in 2017. He is currently pursuing his Ph.D. degree in South China University of Technology, Guangzhou, China. His current research interests include alpha matting, image processing and machine learning.

Han Huang (Senior Member, IEEE) received the B.Man. degree in information management and information system from the School of Mathematics, South China University of Technology (SCUT), Guangzhou, China, in 2003, and the Ph.D. degree in computer science from SCUT in 2008. His research interests include theoretical foundation and application of evolutionary computation and micro-computation. Prof.Huang is a Distinguished Member of CCF.

Yihui Liang received the B.S. degree in digital media technology from Xi'an University of Technology, China, in 2012, the M. Eng. degree and the Ph.D. degree in software engineering from South China University of Technology, China, in 2015 and 2019, respectively. He is currently an associate professor with the School of Computer Science, University of Electronic Science and Technology of China, Zhongshan Institute. His current research interests include alpha matting and image processing.

Zeyang Liu received the B.S. degree in electronic science and technology from China University of Geosciences, Wuhan, China, in 2017, and the M.S. degree in software engineering from South China University of Technology, Guangzhou, China, in 2021. His main research focuses on artificial intelligence.

Zhifeng Hao received the B.Sc. degree in mathematics from Sun Yat-sen University, Guangzhou, China, in 1990, and the Ph.D. degree in mathematics from Nanjing University, Nanjing, China, in 1995. He is currently a Professor with the College of Science, Shantou University, Shantou, Guangdong, China. His research interests include various aspects of algebra, machine learning, data mining, and evolutionary algorithms.