

PAPER

A Mixture Model for Image Boundary Detection Fusion

Yinghui ZHANG^{†a)}, Nonmember, Hongjun WANG^{††b)}, Member, Hengxue ZHOU^{††c)},
and Ping DENG^{††d)}, Nonmembers

SUMMARY Image boundary detection or image segmentation is an important step in image analysis. However, choosing appropriate parameters for boundary detection algorithms is necessary to achieve good boundary detection results. Image boundary detection fusion with unsupervised parameters can output a final consensus boundary, which is generally better than using unsupervised or supervised image boundary detection algorithms. In this study, we theoretically examine why image boundary detection fusion can work well and we propose a mixture model for image boundary detection fusion (MMIBDF) to achieve good consensus segmentation in an unsupervised manner. All of the segmentation algorithms are treated as new features and the segmentation results obtained by the algorithms are the values of the new features. The MMIBDF is designed to sample the boundary according to a discrete distribution. We present an inference method for MMIBDF and describe the corresponding algorithm in detail. Extensive empirical results demonstrate that MMIBDF significantly outperforms other image boundary detection fusion algorithms and the base image boundary detection algorithms according to most performance indices.

key words: expectation maximization, image boundary detection fusion, image segmentation, mixture model

1. Introduction

Image boundary detection or image segmentation is the main and most important method for extracting information from an image. In this method, an image is divided into several parts where each part contains a particular feature [12]. It is easy to find interesting features in the source image based on these outputs. Image segmentation is applied widely in image retrieval, image recognition, and other fields, where a complex image is segmented into simple and readily analyzed parts. The source image only contains insufficient and uncertain information about the object, whereas global information is required. Image boundary detection and fusion are used widely in areas such as remote sensing, automatic control, medical imaging, and computer graphics. Image segmentation or image boundary detection is a key step in image analysis, and it plays important roles

in object recognition, occlusion boundary estimation, image compression, image editing, and image retrieval. Many image segmentation and image boundary detection methods require appropriate parameters to achieve good results. In general, unsupervised approaches are not suitable for effectively selecting the best parameters [1], [24], [37], so the parameters are typically selected by using supervised learning methods or manual approaches.

Image boundary detection fusion is an alternative method for obtaining good results without selecting any parameters, where it can be regarded as a symbolic level of image fusion that uses several boundary detection algorithms or an algorithm with different unsupervised parameters to process an image and obtain several boundaries. The boundaries are then combined in an unsupervised manner to obtain a consensus boundary. Several previous studies [6], [9], [44], [45], [48] have addressed this problem, but three fundamental drawbacks affect existing image boundary detection fusion methods, as follows.

- (i) There have been no theoretical analyses of image boundary detection fusion and it is still unclear why image boundary detection fusion works well.
- (ii) High volumes of computations are required because images normally comprise a large number of pixels. For example, a rather small image of 400×400 pixels (each pixel is regarded as an object) would yield a dataset of 160000 objects [9]. In addition, graph-based and similarity or dissimilarity matrix-based approaches [9] are used for fusion, where a 160000×160000 matrix is produced for computation.
- (iii) Most of the existing algorithms for image boundary detection fusion are designed in a discriminate manner and there is no generative model for solving the fusion problem. Thus, we must test the generative model to solve the problem and determine whether it works.

To address these problems, in this study, we propose a mixture model for image boundary detection fusion (MMIBDF) to solve the image boundary detection fusion problem. MMIBDF addresses the image boundary detection fusion problem by using a mixture model to effectively maintain a multinomial distribution over all possible consensus boundaries. MMIBDF treats the boundary result obtained by an image boundary detection algorithm for each object as a feature vector with discrete feature values, and it learns a mixture model from this feature representation so its com-

Manuscript received September 27, 2017.

Manuscript revised January 5, 2018.

Manuscript publicized January 18, 2018.

[†]The author is with Software Center, Northeastern University, Shenyang 110819, China.

^{††}The authors are with School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China.

a) E-mail: scuky@163.com

b) E-mail: wanghongjun@swjtu.edu.cn

c) E-mail: szhou0000@my.swjtu.edu.cn

d) E-mail: 785417049@my.swjtu.edu.cn

DOI: 10.1587/transinf.2017EDP7314

plexity is relatively low. Our extensive empirical evaluations demonstrated that MMIBDF outperformed other image boundary detection fusion algorithms as well as the base image boundary detection algorithms with different parameters in terms of the most popular performance indices. The present study makes three main contributions, as follows.

- (i) First, we demonstrate why image boundary detection fusion is needed. Based on information theory, we proved that image boundary detection fusion can outperform a single algorithm for boundary detection and obtain better results.
- (ii) Second, our proposed algorithm for image boundary detection fusion is totally unsupervised, where the results obtained by base boundary detection are viewed as new features, which can simplify the image boundary detection fusion process.
- (iii) Third, the result obtained by each base image boundary detection algorithm in the model can be regarded as a multinomial distribution, and a multinomial mixture model is designed for image boundary detection fusion. We propose a generative model to explore the image boundary detection fusion problem.

The remainder of this paper is organized as follows. In Sect. 2, we provide a detailed overview of related research. In Sect. 3, we introduce the object function for image boundary detection fusion based on information theory. We then prove that under certain conditions, the image boundary detection fusion results are very close to or equal to the ground truth. In Sect. 4, we explain the proposed MMIBDF method in detail, including the inference method and a description of the algorithm. Our experimental results are presented in Sect. 5. Finally, we give our conclusions in Sect. 6.

2. Related Work

Many image segmentation or boundary detection algorithms have been proposed in the last 40 years. Since the 1970s, many studies have addressed the problem of image segmentation or image boundary detection, and great progress has been achieved. In general, image segmentation methods can be categorized according to five classes: (1) Threshold methods [7], [8]; (2) Edge detection methods [10], [25]; (3) Statistical-based methods [15], [17]; (4) Region-based segmentation methods [38]; and (5) Methods based on specialized information [27].

Image fusion [5] aims to reduce uncertain information and minimize redundancy, enhance the reliability, and maximize the relevant information for an objective task. The same source image can yield different fusion results because the objectives of fusion tasks and the relevant information are different [14].

Recently, several highly cited studies have investigated the theory and practice of image fusion (e.g., [4], [34], [39]). The discrete wavelet transform is a method used for fusing images in many applications [20], and the dual-tree complex wavelet transform was proposed more recently [18]. Feature

level-based fusion algorithms first divide the image into several sub-images, before fusing them according to the characteristics of the sub-images [19], [32], [49]. Feature level fusion is not sensitive to noise [33]. In multi-sensor systems, each sensor camera has specific parameters so image registration is needed. Image registration is also based on image fusion [13], [50].

Lucas et al. [9] constructed a new framework that uses ordinary clustering and ensemble algorithms to solve the image boundary detection fusion problem. Vian et al. regarded the fusion problem as a feature selection problem, where they utilized a selection criterion based on mutual information. Starting with the object regions detected roughly by one sensor, the technique proposed by Ren [43] aims to extract relevant information from another sensor in order to complete the object segmentation process. Furthermore, the fusion problem decision map is a basic concept where the decision map decides the information that will be extracted from each specific region. Li et al. [21] proposed an algorithm that uses the classical k -means algorithm to segment input images, before employing a genetic algorithm to fuse them, where this method has been used for medical tumor detection [35].

Recently, image boundary detection fusion (combination) has attracted much attention because of its various advantages, where it aims to output a final consensus segmentation result based on a set of different segmentation results but without requiring supervision. Different combination methods [6], [44], [48] that consider the size of the data or the structure of the pattern's lattice have been designed to deal with the image segmentation combination problem. Xu et al. [45] employed the mean shift algorithm to obtain an initial boundary contour for an object and the Canny edge detection algorithm to extract salient edges, before integrating the results. A segmentation ensemble method based on the weighted partition consensus [42] was proposed based on kernel methods [41], where the main limitation that affected this approach was an image with large number of pixels. Therefore, a symmetric matrix containing the number of pixels in the image was used in the worst case for his algorithm, which requires large amounts of computations. A random walker approach [44] was also proposed to combine multiple segmentation results to obtain an approximation of the generalized median segmentation result.

3. Image Boundary Detection Fusion

In the following, we describe the problem of image boundary detection fusion in detail and theoretically prove that the image boundary detection fusion results can be better than the base image boundaries. The notations used in this study are summarized in Table 1.

3.1 Problem Definition

Boundary detection fusion aims to combine several boundaries into a consensus boundary that represents the opti-

Table 1 Notations

Symbol	Explanation
M	Number of boundaries
N	Number of pixels
P	Set of image pixels
p	Pixel in an image
X, λ	Boundaries obtained by different boundary detection algorithms
s	Base boundary detection algorithm
K	Number of values in all boundaries
K(j)	Number of values in the j th boundary
π	Mixing coefficients
Θ	Parameter for a mixture model
b	Variable boundary

mal boundary. In general, different boundary detection algorithms (or one boundary detection algorithm with different parameters) are first used to generate a set of different boundaries. An image boundary detection fusion algorithm is then used to obtain a consensus boundary.

Given N pixels $P = \{\mathbf{p}_i, [i]_1^N\}$ in an original image and M image boundary detection algorithms (or one boundary detection algorithm with M parameters) $S = \{s_j, [j]_1^M\}$, we obtain M boundary results, with one from each algorithm. If λ_{ij} denotes the boundary assigned to \mathbf{p}_i by s_j , then the boundary detection algorithm s_j yields a segmentation of the entire original pixels given by $\lambda_j = \{\lambda_{ij}, [i]_1^N\} = \{s_j(\mathbf{p}_i), [j]_1^M\}$. Image boundary detection fusion can be solved by finding a median partition λ^* from $\{\lambda_j, [j]_1^M\}$. In this study, the function $d(\lambda^*, \lambda_j)$ is the symmetric difference distance metric and we aim to find a set partition λ^* such that

$$\lambda^* = \arg_{\lambda} \min \sum_{j=1}^M d(\lambda_j, \lambda). \quad (1)$$

An optimal median boundary can be found by solving this optimization problem, which is NP complete [3]. An exhaustive search approach for the problem has computational complexity of $O(K^N/M!)$ for $N \geq 2$ and $M \geq 2$. Exhaustive search to solve Eq. (1) is computationally expensive because there are $K^N/K!$ possible subsets. However, some deterministic optimization strategies [28], [29] can solve this problem.

3.2 On Image Boundary Detection Fusion

Next, we explain the motivation for boundary detection fusion in detail. We also show that image boundary detection fusion is more robust and stable than any single boundary detection algorithm, where the results obtained can be extremely close to the ground truth. We assume that the results obtained by any boundary detection algorithms are better than random guesses and that they are independent of each other. The majority voting algorithm is used for image boundary detection fusion. First, we show that Eq. (2) is valid,

$$\lim_{M \rightarrow \infty} p_c = 1, p_c > \frac{1}{K}, c = (1, 2, \dots, M), \quad (2)$$

where p_c is the accuracy of boundary detection fusion, p_c is also the accuracy of any single boundary detection algorithm, and M is the number of boundary detection algorithms, where each p_c is independent of others. It seems that $M \rightarrow \infty$ is unacceptable but in a real situation, it is acceptable for M to range from 5–100 and a better result can be obtained. For example, if we have five completely independent boundary detection algorithms, then majority voting is used for boundary detection fusion. If the accuracy is 0.7 for each algorithm, then $10 \times (0.7^3) \times (0.3^2) + 5 \times (0.7^4) \times (0.3) + (0.7^5) = 0.837$ is the majority vote fusion accuracy. If 101 boundary detection algorithms are available, then the majority vote fusion accuracy is 0.999. Each partition of the pixel status (i.e., boundary or not) matches the ground truth with a probability of h ; otherwise, $h > \forall g_i$. Let Y_1 be the number that match the ground truth and Y_i is the number that do not match the ground truth. For M independent segmentations, the random variables Y_1 and Y_i are subject to the multinomial distribution: $Y_1 \sim B(h, M)$, $Y_i \sim B(g_i, M)$, where $h > 1/K$ and $\sum_{i=2}^K g_i + h = 1$. Given M independent partitions, the joint probability of random variables Y_1, Y_i is a multinomial [30] according to the probability mass function[†].

$$\begin{aligned} P(Y_1 = M_1, Y_2 = M_2, \dots, Y_K = M_K) \\ = \frac{M!}{M_1! M_2! \dots M_K!} h^{M_1} g^{M_2} \dots g^{M_K}, \end{aligned} \quad (3)$$

where $M = M_1 + M_2 + \dots + M_K$.

The probability p_c of boundary fusion based on M partitions using majority voting is

$$\begin{aligned} p_c &= P(Y_1 > Y_2, Y_1 = Y_3, \dots, Y_1 = Y_K) \\ &= 1 - P(Y_1 \leq Y_i) \end{aligned} \quad (4)$$

$\forall i \geq 2, P(Y_1 \leq Y_i) \rightarrow 0$ and $M \rightarrow \infty$,

$$p_c \geq 1 - \sum_{i=2}^K P(Y_1 \leq Y_i). \quad (5)$$

Let $Z_i = \frac{Y_1 - Y_i}{M}$. $P(Y_1 \leq Y_i)$ can be rewritten as $P(Z_i \leq 0)$. The expected value and variance of Z_i can be obtained from Eqs. (6)–(7):

[†]https://en.wikipedia.org/wiki/Multinomial_distribution.

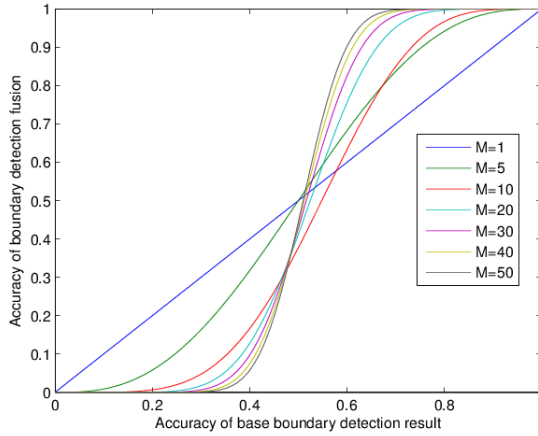


Fig. 1 Relationship between the accuracy of image boundary detection fusion and the accuracy of the boundary detection algorithm.

$$\begin{aligned} E[Z_i] &= E\left[\frac{Y_1 - Y_i}{M}\right] = \frac{1}{M}(E[Y_1] - E[Y_i]) \\ &= \frac{1}{N}(Mh - Mg) = h - g, \end{aligned} \quad (6)$$

$$\begin{aligned} \text{Var}[Z_i] &= \text{Var}\left[\frac{1}{M}(Y_1 - Y_i)\right] \\ &= \frac{1}{M^2}(\text{Var}[Y_1] + \text{Var}[Y_i] - 2\text{Cov}[Y_1, Y_i]) \\ &= \frac{1}{M^2}[Mh(1-h) + Mg(1-g) - 2Mhg] \\ &= \frac{1}{M}[h(1-h) + g(1-g) - 2hg]. \end{aligned} \quad (7)$$

As a result, we have

$$\lim_{M \rightarrow \infty} p_c = \lim_{M \rightarrow \infty} P(Y_1 > Y_i) = 1. \quad (8)$$

From Eq. (8), as the value of M increases, the image boundary detection fusion result becomes closer to the ground truth. According to the theoretical analysis, the accuracy of image boundary detection fusion is given by Eq. (9),

$$p_c = \sum_{i=\lceil \frac{M+1}{2} \rceil}^M \frac{M!}{2!(M-i)!} h^2 (1-h)^{M-2}. \quad (9)$$

A graphical representation of Eq. (8) where $M = \{1, 5, 10, 20, 30, 40, 50\}$ is shown in Fig. 1, which clearly indicates that the accuracy of image boundary detection fusion increases with M . In addition, the accuracy of image boundary detection converges to 1 as the number of boundary detection algorithms M increases.

We have proved that image boundary detection fusion is much better than image boundary detection, where its computational complexity is $O(NM!/K!(M-K)!)$ if direct and greedy optimization approaches are used to solve the image boundary detection problem. Thus, in the next section, we propose an effective image boundary detection fusion approach based on the mixture model.

4. MMIBDF

The base boundary detection results are transformed from rows into columns and we propose an image boundary detection fusion method based on a finite mixture model of the probability of the consensus boundary.

The results obtained by M boundary detection algorithms can be stacked together to form an $(N \times M)$ matrix where the j^{th} column is λ_j . This matrix can be viewed from another perspective where each row \mathbf{x}_i of the matrix, *i.e.*, all of the boundary detection results for \mathbf{p}_i , gives a new feature vector representation comprising pixels \mathbf{p}_i . In particular, $\mathbf{x}_i = \{x_{ij}, [j]_1^M\} = \{s_j(\mathbf{p}_i), [j]_1^M\}$. Given the boundary detection matrix B , solving the image boundary detection fusion problem involves combining the M boundary detection results for N pixels to generate a consensus boundary, which might be more accurate, robust, and stable than the individual boundary detection algorithm.

When using the MMIBDF, the main assumption is that x_i is modeled as a random variable drawn from a probability distribution described as a mixture of multivariate component densities:

$$p(x_i|\Theta) = \sum_{k=1}^K \pi_k P_k(x_i|\theta_k) \quad (10)$$

where each component is parameterized by θ_k . The K components of the mixture are identified with the boundary detection algorithms s_j . The mixing π_k corresponds to the prior probabilities of the boundary or not. In this model, a data point x_i is generated in two steps by drawing a component according to the probability mass function π_k and then sampling a point from the distribution $P_m(x|\theta_m)$. All of the data points $X = \{x_i, i_1^N\}$ are assumed to be independent and identically distributed. Thus, for given the data set X , the log-likelihood function for the parameters Θ is represented as:

$$\log L(\Theta|Y) = \log \prod_{i=1}^N P(x_i|\Theta) = \sum_{i=1}^N \log \sum_{k=1}^K \pi_k P_k(x_i|\theta_k). \quad (11)$$

The objective of image boundary detection fusion is now formulated as a maximum-likelihood estimation problem. To find the best fitting mixture density for the given data X , the likelihood function is maximized as: $\Theta^* = \arg \max_{\Theta} \log L(\Theta|X)$. From Eq. (11), the most important term is the model of component-conditional densities $P_k(x|\theta_k)$. To make the problem more tractable, a conditional independence assumption is made regarding the components of the vector x_i . The conditional probability of x_i can be represented as the following product:

$$P_k(x_i|\theta_k) = \prod_{j=1}^M P_k^j(x_{ij}|\theta_k^j).$$

For the mixture model, this is the choice of a probability density $P_m^j(x_{ij}|\theta_m^j)$ for the components of the vectors x_i . The

variable $x_{i,j}$ comprises the nominal values from the results obtained by the boundary detection algorithm s_j , so it is natural to view them as the outcome of a multinomial trial: $P_k^j(x_{ij}|\theta_k^j) = \prod_{k=1}^{K(j)} \theta_{jk}(k)^{\delta(x_{ij},k)}$, where k is a mixture parameter. In general, the maximum-likelihood problem in Eq. (11) cannot be solved in a closed form when all of the parameters Θ are unknown. However, the likelihood function can be optimized by the expectation maximization (EM) algorithm.

Thus, we assume that the hidden variable $Z = \{z_1, z_i, \dots, z_M\}$ exists and that the likelihood of the complete data is (X, Z) . If the value of z_i is known, then we can immediately determine which of the M mixture components is used to generate the point x_i . The detailed derivation of the EM solution to the mixture model with multivariate Bernoulli components leads to the equations for the expectation (E)-step and maximization (M)-step, which are repeated in each iteration of the algorithm:

$$E[z_{ik}] = \frac{\pi_k^t \prod_{j=1}^M \prod_{k=1}^K (\theta_{jk}^t(k))^{\delta(x_{ij},k)}}{\sum_{n=1}^M \pi_n^t \prod_{j=1}^M \prod_{k=1}^K (\theta_{jk}^t(k))^{\delta(x_{ij},k)}}, \quad (12)$$

$$\pi_k = \frac{\sum_{i=1}^N E[z_{ik}]}{\sum_{i=1}^N \sum_{k=1}^K E[z_{ik}]}, \quad (13)$$

$$\theta_{jk}(k) = \frac{\sum_{i=1}^N \delta(x_{ij},k) E[z_{ik}]}{\sum_{i=1}^N \sum_{k=1}^K \delta(x_{ij},k) E[z_{ik}]}. \quad (14)$$

The solution to the image boundary detection fusion problem is obtained simply by inspecting the expected values of the variables $E[z_{ik}]$ because $E[z_{ik}]$ represents the probability that the pattern x_i is generated by the mixture component.

The proposed model can also combine with coarse boundaries, where we view coarse boundaries as missing labels (boundaries). It is also possible to apply the EM algorithm in the case of missing data. For coarse boundaries, each vector x_i in X can be split into observed and missing components $x_i = (x_i^o, x_i^u)$. Incorporating missing data requires a slight modification of the computation of the E- and M-steps. First, the expected values $E[z_{ik}|x_i^o, \Theta^t]$ are inferred based on the observed components of vector x_i , i.e., the products in Eq. (12) are taken over known labels: $\prod_{j=1}^M \rightarrow \prod_{j:x^o}$. In addition, we must compute the expected values $E[z_{ik}|x_i^k|x_i^o, \Theta^t]$ and substitute them, as well as $E[z_{ik}|x_i^k|x_i^o, \Theta^t]$, into the M-step for re-estimating the parameters $\theta_{jk}(k)$.

The image boundary detection fusion algorithm based on the mixture model is summarized as follows. Starting from an initial guess $\Theta = \{\pi_0, \theta_0\}$, the EM algorithm alternates between two E- and M-steps until convergence is reached.

MMIBDF algorithms 1:

Input: The final K and image boundary detection data, $\{X = x_{ij}, [i]_1^N, [j]_1^M\}$, where x are the results obtained by all the boundary detection algorithms.

Output: $s_c(x_i)$, the final consensus segmentation result.

-
- (i) E-Step: Given $(\pi^{(t-1)}, \theta^{(t-1)})$, for each x_i , find the best expected values $E[z_{ik}]$. If there are missing data, then find the best expected values $\{E[z_{ik}|x_i^u]\}$.
 - (ii) M-Step: Maximize and re-estimate parameters $\{\theta_{jk}(k)\}$.
 - (iii) $s_c(x_i) = \max(z_i)$, find the index of the component of z_i with the largest expected value.
-

The computational complexity of the proposed EM procedure is $o(TNM + TKN^2)$, where T is the number of iterations. The algorithm produces a monotonically increasing likelihood by repeating the E- and M-steps alternately until a local maximum or global maximum is approached. The proposed algorithm is much more effective compared with direct and greedy optimization approaches. We did not use the greedy method in the experimental evaluations because it requires an excessive amount of time.

5. Empirical Study

We evaluated the results obtained by MMIBDF based on real images from the Berkeley segmentation database[†], which also contains many boundary detection results produced by popular algorithms [1], [37] and the ground truth results obtained by human subjects [46]. All of the experiments were conducted using Matlab R2010a as the platform on a desktop machine (with two Intel Xeon processors at 2.4 GHz and 48 GB RAM). The experimental setups employed are described in the following.

- (i) We used the `yu` [47], `xren_gray` [36], `ren_nips2012_gray` [37], `gPb_gray` [22], and `gPb-ucm` [1] algorithms to obtain base boundaries (if the results were expressed in segmentation format, the segmentation format was changed into boundary format). Thus, five base boundaries were obtained for an image and they were used as the inputs for the fusion algorithms.
- (ii) The MMIBDF, QMlc [40], HGPA [9], MCLA [9], and vote [2] algorithms were used for boundary detection fusion.

For the evaluations, we used the probabilistic rand index (PRI) [31], variation of information (VoI) [26], global consistency error (GCE) [23], and boundary displacement error (BDE) [11] as performance measures. The results are shown in Table 2, where the ground truths for BSDS300 were downloaded from the web data set^{††} and its subdirectory. The results obtained by `yu`, `BEL`, and other methods were obtained from the web data set^{†††}.

[†]<http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/bench/html/images.html>

^{††}http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/bench/gray/gPb-ucm_gray/main.html

^{†††}<http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/bench/html/algorithms.html>

Table 2 Average performance based on the test images from BSDS300 and the ground truth results. Bold indicates the best result obtained by all of the algorithms.

	PRI	GCE	VoI	BDE
yu	0.7608	0.0384	0.7230	18.7830
xren_gray	0.7858	0.0365	1.7092	13.6744
ren_nips2012_gray	0.7547	2.1214	2.1935	11.4922
gPb_gray	0.7800	0.0362	1.8798	12.4925
gPb-ucm	0.8102	0.0386	1.6476	11.7604
QMlc	0.3940	0.0158	7.5016	13.9031
HGPA	0.5843	0.0375	2.4491	12.7070
MCLA	0.8317	0.0382	0.5581	12.4757
vote	0.6799	0.0373	2.2520	11.3785
MMIBDF	0.8386	0.0363	0.5418	12.7213

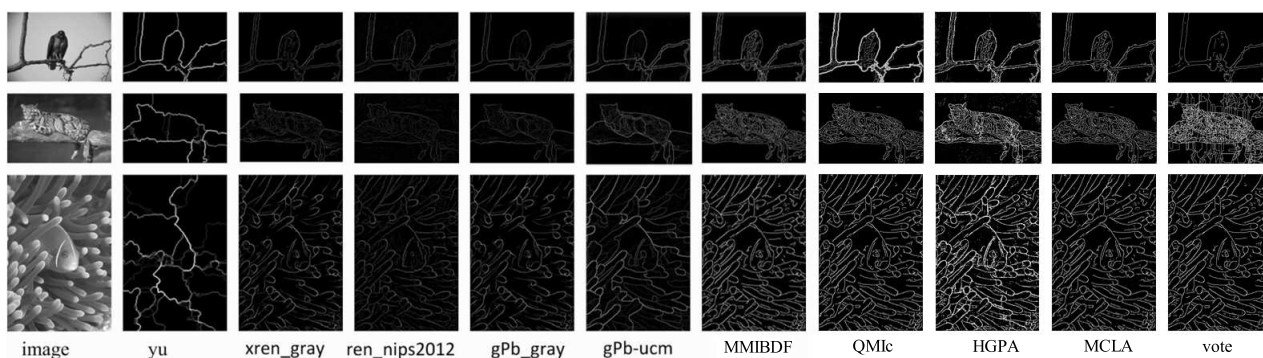


Fig. 2 Original image, base boundary detection results, and fusion results.

- (i) PRI counts the overall fraction of pairs of pixels with labels that are consistent between two segmentations or the segmentation and the ground truth. PRI ranges between $[0, 1]$, where a higher PRI value denotes better performance. Table 2 shows clearly that MMIBDF performed significantly better than all of the other fusion algorithms and popular segmentation algorithms.
- (ii) VoI is the distance between two segmentations as the average conditional entropy of one segmentation given the other, and it roughly measures the amount of randomness in one segmentation. VoI ranges between $[0, \infty)$, where a lower VoI value represents a better segmentation result. Table 2 shows clearly that MMIBDF performed significantly better than all the other fusion algorithms and popular segmentation algorithms.
- (iii) GCE measures the extent to which one segmentation can be viewed as a refinement of the ground truth. The related segmentations in GCE are considered to be consistent because it is a good method for representing the same natural image segmented at different scales. GCE ranges between $[0, \infty)$, where a lower value is better. Table 2 shows clearly that MMIBDF obtained the third best result.
- (iv) BDE measures the average displacement error of the boundary pixels between two segmentations. BDE ranges between $[0, \infty)$, where the boundary performance is better when the BDE value is lower. Table 2 shows that the results obtained by MMIBDF were not the best and its performance was ranked 7_{th} among the 10 algorithms. The BDE value obtained for the proposed model was slightly inferior because the proposed

model sometimes obtains a local optimal result, which is amplified by the BDE measure.

Thus, in general, MMIBDF performed better than all of the other algorithms and it obtained the best results in terms of most of the indices, such as PRI and BDE. MMIBDF also had the third best performance in terms of VoI. However, the results obtained by MMIBDF were not good according to the BDE value. The experimental results demonstrate that the proposed model performs well.

Moreover, the results obtained using MMIBDF, MCLA, and Bayesian image boundary fusion (BISF) were similar in terms of the PRI, VoI, GCE, and BDE indices. MCLA is a good algorithm for correspondences when there is little noise and low diversity among base boundaries, and thus it obtained similar results to MMIBDF, although they were lower than produced using MMIBDF.

We also performed visual inspections in the evaluation and the results are shown in Fig. 2. The boundaries were generated by the yu, xren_gray, ren_nips2012_gray, gPb_gray, gPb-ucm, and MMIBDF algorithms. It is difficult to see the differences between the results and the best, but the proposed model clearly performed well.

6. Conclusions

In this study, we formalized image boundary detection fusion as a combinatorial optimization problem and MMIBDF was proposed to obtain a good consensus boundary. In our method, we treat the boundary detection algorithms (or the same algorithm with different parameters) as new features

and the boundary results as the values of the new features, which simplifies the image boundary detection fusion problem solving process. We designed a generative model for MMIBDF to sample the boundary according to a discrete distribution. The inference method for MMIBDF and the corresponding algorithm were explained in detail. Finally, real images from the Berkeley segmentation database were used in experimental evaluations and the results demonstrated that MMIBDF performed significantly better than other image boundary detection fusion algorithms and base image boundary detection algorithms. In future research, we will test our method with noisy images and study how the MMIBDF can converge to a global optimum, while a Gibbs sampling method for inference in the proposed model will also be studied.

Acknowledgments

This study was partially supported by NSFC (No. 61262058), the Fundamental Research Funds for the Central Universities (No. A0920502051515-12), and Online Education Research Center of the Ministry of Education Online Education Research Fund (No. 2016YB158). We thank the anonymous reviewers for their helpful comments and suggestions.

References

- [1] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.33, no.5, pp.898–916, 2011.
- [2] H.G. Ayad and M.S. Kamel, "On voting-based consensus of cluster ensembles," *Pattern Recognition*, vol.43, no.5, pp.1943–1953, 2010.
- [3] J.-P. Barthélemy and B. Leclerc, "The median procedure for partitions," *DIMACS series in discrete mathematics and theoretical computer science*, vol.19, pp.3–34, American Mathematical Society, 1995.
- [4] R.S. Blum and Z. Liu, *Multi-sensor image fusion and its applications*, CRC press, 2005.
- [5] P.J. Burt and E.H. Adelson, "Merging images through pattern decomposition," *SPIE*, vol.575, pp.173–181, 1985.
- [6] K. Cho and P. Meer, "Image segmentation from consensus information," *Computer Vision and Image Understanding*, vol.68, no.1, pp.72–89, 1997.
- [7] C.M. Christoudias, B. Georgescu, and P. Meer, "Synergism in low level vision," *Proceedings. 16th IEEE International Conference on Pattern Recognition*, vol.4, pp.150–155, 2002.
- [8] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.5, pp.603–619, 2002.
- [9] L. Franek, D.D. Abdala, S. Vega-Pons, and X. Jiang, "Image segmentation fusion using general ensemble clustering methods," *Computer Vision—ACCV 2010*, vol.6495, pp.373–384, Springer, 2011.
- [10] W.T. Freeman and E.H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.13, no.9, pp.891–906, 1991.
- [11] J. Freixenet, X. Muñoz, D. Raba, J. Martí, and X. Cufi, "Yet another survey on image segmentation: Region and boundary information integration," *Proceedings of European Conference on Computer Vision*, vol.2352, pp.408–422, 2002.
- [12] B. Georgescu, I. Shimshoni, and P. Meer, "Mean shift based clustering in high dimensions: A texture classification example," *Proc. Ninth IEEE International Conference on Computer Vision*, pp.456–463, 2003.
- [13] A.A. Goshtasby, "Fusion of multifocus images to maximize image information," *Defense and Security Symposium*, pp.62290L–62290L-10, International Society for Optics and Photonics, 2006.
- [14] A.A. Goshtasby and S. Nikolov, "Image fusion advances in the state of the art," *Information Fusion*, vol.8, no.2, pp.114–118, April 2007.
- [15] L. Gupta and T. Sortrakul, "A Gaussian-mixture-based image segmentation algorithm," *Pattern Recognition*, vol.31, no.3, pp.315–325, Oct. 1998.
- [16] J. Kim, J.W. Fisher, A. Yezzi, M. Cetin, and A.S. Willsky, *Nonparametric statistical methods for image segmentation and shape analysis*, PhD thesis, Massachusetts Institute of Technology, pp.797–800, 2005.
- [17] J. Kim, J.W. Fisher, A. Yezzi, M. Çetin, and A.S. Willsky, "A non-parametric statistical method for image segmentation using information theory and curve evolution," *IEEE Trans. Image Process.*, vol.14, no.10, pp.1486–1502, 2005.
- [18] N. Kingsbury, *Image Processing with Complex Wavelets*, Oxford University Press, 1999.
- [19] J.J. Lewis, R.J.O. Callaghan, S.G. Nikolov, D.R. Bull, and C.N. Canagarajah, "Region-based image fusion using complex wavelets," *Seventh International Conference on Information Fusion*, vol.1, pp.555–562, 2004.
- [20] H. Li, B.S. Manjunath, and S.K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol.57, no.3, pp.235–245, 1995.
- [21] S. Li and B. Yang, "Multifocus image fusion using region segmentation and spatial frequency," *Image and Vision Computing*, vol.26, no.7, pp.971–979, 2008.
- [22] M. Maire, P. Arbeláez, C. Fowlkes, and J. Malik, "Using contours to detect and localize junctions in natural images," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1–8, 2008.
- [23] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," *Proceedings of IEEE International Conference on Computer Vision*, pp.416–423, 2001.
- [24] D.R. Martin, C.C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.26, no.5, pp.530–549, 2004.
- [25] L. Matalas, R. Benjamin, and R. Kitney, "An edge detection technique using the facet model and parameterized relaxation labeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.19, no.4, pp.328–341, 1997.
- [26] M. Meilă, "Comparing clusterings: an axiomatic view," *Proceedings of the International Conference on Machine Learning*, pp.577–584, 2005.
- [27] R. Mester and U. Franke, "Statistical model based image segmentation using region growing, contour relaxation and classification," *Proceedings of the SPIE Visual Communications and Image Processing*, pp.616–624, 1988.
- [28] M. Mignotte, "A label field fusion bayesian model and its penalized maximum rand estimator for image segmentation," *IEEE Trans. Image Process.*, vol.19, no.6, pp.1610–1624, 2010.
- [29] M. Mignotte, "A label field fusion model with a variation of information estimator for image segmentation," *Information Fusion*, vol.20, pp.7–20, 2014.
- [30] J.E. Mosimann, "On the compound multinomial distribution, the multivariate β -distribution, and correlations among proportions," *Biometrika*, vol.49, no.1/2, pp.65–82, 1962.
- [31] C. Pantofaru and M. Hebert, "A comparison of image segmentation algorithms," *Tech. Rep. CMU-RI-TR-05-40*, Carnegie Mellon University, vol.24, 2005.
- [32] G. Piella, "A general framework for multiresolution image fusion: from pixels to regions," *Information Fusion*, vol.4, no.4,

- pp.259–280, 2003.
- [33] G. Piella, “A region-based multiresolution image fusion algorithm,” *Proceedings of the Fifth IEEE International Conference on Information Fusion*, vol.2, pp.1557–1564, 2002.
 - [34] P.J. Burt and R.J. Kolczynski, “Enhanced image capture through fusion,” *International Conference on Computer Vision*, pp.173–182, 1993.
 - [35] S.M. Praveena and I. Vennila, “Optimization fusion approach for image segmentation using k-means algorithm,” *International Journal of Computer Applications*, vol.2, no.7, pp.18–25, 2010.
 - [36] X. Ren, “Multi-scale improves boundary detection in natural images,” *European Conference on Computer Vision*, pp.533–545, Springer, 2008.
 - [37] X. Ren and L. Bo, “Discriminatively trained sparse code gradients for contour detection,” *Advances in Neural Information Processing Systems*, pp.584–592, 2012.
 - [38] N. Senthilkumaran and R. Rajesh, “Edge detection techniques for image segmentation—a survey of soft computing approaches,” *International Journal of Recent Trends in Engineering*, vol.1, no.2, pp.250–254, 2009.
 - [39] M.I. Smith and J.P. Heather, “A review of image fusion technology in 2005,” *Defense and Security*, pp.29–45, International Society for Optics and Photonics, 2005.
 - [40] A. Topchy, A.K. Jain, and W. Punch, “Clustering ensembles: Models of consensus and weak partitions,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.27, no.12, pp.1866–1881, 2005.
 - [41] S. Vega-Pons, J. Correa-Morris, and J. Ruiz-Shulcloper, “Weighted partition consensus via kernels,” *Pattern Recognition*, vol.43, no.8, pp.2712–2724, 2010.
 - [42] S. Vega-Pons, X. Jiang, and J. Ruiz-Shulcloper, “Segmentation ensemble via kernels,” 2011 *IEEE Asian Conference on Pattern Recognition*, pp.686–690, 2011.
 - [43] S. Vinay and W.D. James, “Feature-level fusion for object segmentation using mutual information,” *Augmented Vision Perception in Infrared*, pp.295–320, Springer, 2009.
 - [44] P. Wattuya, K. Rothaus, J.-S. Pražni, and X. Jiang, “A random walker based approach to combining multiple segmentations,” 19th *IEEE International Conference on Pattern Recognition*, pp.1–4, 2008.
 - [45] W. Xu, R. Kanawong, Y. Duan, and G. Zhang, “A new information fusion approach for image segmentation,” 18th *IEEE International Conference on Image Processing*, pp.2873–2876, 2011.
 - [46] A.Y. Yang, J. Wright, Y. Ma, and S.S. Sastry, “Unsupervised segmentation of natural images via lossy data compression,” *Computer Vision and Image Understanding*, vol.110, no.2, pp.212–225, 2008.
 - [47] S.X. Yu, “Segmentation induced by scale invariance,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol.1, pp.444–451, 2005.
 - [48] Z. Yu, S. Zhang, H. Wong, and J. Zhang, “Image segmentation based on cluster ensemble,” *Advances in Neural Networks*, vol.4993, pp.894–903, 2007.
 - [49] Z. Zhang and R. Blum, “Region-based image fusion scheme for concealed weapon detection,” *Proceedings of the 31st Annual Conference on Information Sciences and Systems*, pp.168–173, 1997.
 - [50] Z. Zhang and R.S. Blum, “A hybrid image registration technique for a digital camera image fusion application,” *Information Fusion*, vol.2, no.2, pp.135–149, June 2001.



Yinghui Zhang received his Ph.D. degree in computer science from Northeastern University of China in 1998. He is currently a professor and dean at Chengdu Neusoft University. His research interests are cloud computing, data mining, and knowledge discovery engineering. He has published over 40 research papers in journals and conferences, and he is a senior member of the CCF. He has been a reviewer for several academic journals.



Hongjun Wang received his Ph.D. degree in computer science from Sichuan University of China in 2009. He is currently an associate professor in the Key Lab of Cloud Computing and Intelligent Techniques at Southwest Jiaotong University. His research interests are machine learning, data mining, and ensemble learning. He has published over 70 research papers in journals and conferences, and he is a member of the ACM and CCF. He has been a reviewer for several academic journals.



Hengxue Zhou is a master student in the School of Information Science and Technology, Southwest Jiaotong University. He received his bachelor degree from Southwest Jiaotong University in 2016. His research interests include data mining and image processing.



Ping Deng is a master student in the School of Information Science and Technology, Southwest Jiaotong University. She received her bachelor degree from China West Normal University in 2016. Her research interests include ensemble learning and image processing.