

An Efficient Image Co-segmentation Algorithm based on Active Contour and Image Saliency

Zhizhi Zhang¹, Xiabi Liu¹, Nouman Qadeer Soomro²

¹Beijing Laboratory of Intelligent Information Technology, School of Computer, Beijing Institute of Technology, Beijing 100081, China

²Department of Computer Science, DHA Suffa University, Karachi Pakistan

Abstract. Image co-segmentation is the problem of extracting common objects from multiple images and it is a very challenging task. In this paper we try to address the co-segmentation problem by embedding image saliency into active contour model. Active contour is a very famous and effective image segmentation method but performs poor results if applied directly to co-segmentation. Therefore, we can introduce additional information to improve the segmentation results, such as saliency which can show the region of interest. In order to optimize the model, we propose an efficient level-set optimization method based on super-pixels, hierarchical computation and convergence judgment. We evaluated the proposed method on iCoseg and MSRC datasets. Compared with other methods, our method yielded better results and demonstrated the significance of using image saliency in active contour.

1 Introduction

Co-segmentation is a relatively new and fast emerging research field in computer vision. Different from the traditional segmentation, which separates the foreground objects from the background, co-segmentation is defined as the simultaneous segmentation of similar regions from two (or more) images. The necessity of co-segmentation is increasing in recent years as many applications require desired-objects to be carved-out from bulk of images or videos, stored in scattered and haphazard manner. Co-segmentation is difficult due to similar backgrounds and variations in targeted objects.

The existing co-segmentation methods mainly include Markov Random Field (MRF) based, active contour based, clustering based, user interaction based and graph based. MRF based co-segmentation methods include Rother et al. [1], Mukherjee et al. [2], Hochbaum et al. [3] and Vicente et al. [4]. Rother et al. [1] is the first work of co-segmentation. They used the l_1 -norm to measure the foreground consistency, and used the Trust Region Graph Cut (TRGC) algorithm to optimize the co-segmentation model. After Rother et al. [1], several other co-segmentation methods were proposed by improving the model construction and optimization. Vicente et al. [4] summarized the previous works [1-3] and presented a new MRF model based on Boykov-Jolly model. To optimize the proposed model, the Dual Decomposition (DD) optimization method was used. Co-segmentation methods based on clustering mainly include Joulin et al. [5]. They proposed a co-segmentation method based on discriminative clustering. In their paper, normalized cuts and kernel methods are used in a discriminative

clustering framework: the goal is to assign foreground/background labels jointly to all images, so that a supervised classifier trained with these labels leads to maximal separation of the two classes [5]. User interaction based methods mainly include Batra et al. [6]. They proposed co-segmentation method called iCoseg based on a few user interactions. iCoseg can intelligently recommend regions to scribble on, and users following these recommendations can achieve good quality cutouts with significantly less time and energy than exhaustively examining all cutouts [6]. Vicente et al. [7] introduced graph based method. They presented the concept of object co-segmentation, which focused on segmenting “objects” rather than “stuff”. In their work, they firstly got the candidate object-like segmentations and then described the relationship between the regions by the fully connected graph. Finally, Loop Belief Propagation (LBP) was used to obtain results. In 2012, Meng et al. [8] proposed a co-segmentation model based on active contour, and later in 2013 Meng et al. [9] improved the co-segmentation method. The ideas of former and later works were somehow the same. Based on these two closely related works, energy function was established on foreground similarity among the images and background consistency in each image. Then level-set formulation and the optimization are illustrated.

This paper presents a co-segmentation method by introducing image saliency into well-known active contour based segmentation method and we called ACIS. According to active contour schema, we design an energy function, in which the image saliency is considered for giving additional information of foreground. The energy function is minimized by using the level-set strategy. In

order to speed up the computation, we propose an efficient level-set optimization based on hierarchical segmentation, super-pixel and better termination. We conducted the experiments on widely used iCoseg image database and MSRC image database. Experimental results demonstrated the effectiveness of the proposed co-segmentation algorithm. The contributions of this paper are summarized as follows:

- A new image co-segmentation model is constructed by introducing the saliency information into the active contour framework. Saliency information can bring useful foreground features to achieve better performance.
- An efficient algorithm is developed through combining super-pixels, hierarchical computation and convergence judgment for solving the co-segmentation problem based on the model above.

The rest of this paper is organized as follows. Section 2 introduces the proposed co-segmentation model based on active contour and image saliency. Section 3 describes our optimization algorithm. Experiments and results are discussed in Section 4. We conclude the paper in Section 5.

2 The proposed co-segmentation model

As described in Section 1, the goal of the co-segmentation is extracting common objects from a group of images. Active contour is a classical and effective method for image segmentation. It is possible to apply it on co-segmentation task. Our proposed co-segmentation model considers not only foreground consistency but also background consistency, moreover, not only consistency within a single image but also consistency between images. Besides, we introduced the saliency information into the active contour framework to achieve better performance. According to the active contour algorithm, given original images, each image should have initial contour, and each contour keeps changing according to the energy function until the energy is minimized. In this paper, the initial contour is a rectangle without artificial participation because we observed the initial contour has little effect on the final results in the experiments, and the size of rectangle is determined by the parameters.

Let $I = \{I_1, I_2, \dots, I_n\}$ be a group of images, suppose C_k is the initial contour of I_k . Based on the initial contour set $C^0 = \{C_1, C_2, \dots, C_n\}$, the evolution of the contour curve is performed to achieve common object segmentation. Let ω_{int}^k be the inner region of the k -th image curve, and ω_{ext}^k be the outer region of the k -th image curve. So the goal of the co-segmentation is to get the curve when the energy is minimized:

$$C_k^* = \arg \min_{C_k} E(C_k), i = 1, \dots, n. \quad (1)$$

For a given curve C_k , according to the active contour Chan-Vese (C-V) model [10], our energy function $E(C_k)$ of the curve is expressed as:

$$E(C_k) = E_c^k - \lambda_{sf} E_{sf}^k - \lambda_{sb} E_{sb}^k - \lambda_f E_f^k - \lambda_b E_b^k - \lambda_{co} E_{co}^k, \quad (2)$$

where λ_{sf} , λ_{sb} , λ_f , λ_b , and λ_{co} are weights for each item. The details of each term in Eq. 2 are described as follows.

The regularization term of the curve: E_c^k :

Regularizing terms mainly describe the attribute of curve, including internal curve area and curve length:

$$E_c^k = \mu \cdot Area(\omega_{int}^k) + \nu \cdot Length(C_k), \quad (3)$$

where $Area(\omega_{int}^k)$ refers to curve area, $Length(C_k)$ refers to the curve length, and μ and ν represent the weights for two factors.

The energy of foreground consistency of the k -th image: E_{sf}^k :

The foreground consistency of a single image is used to guarantee the consistency of the foreground pixels. It also ensures the differentiation of foreground from background. When the pixels in the curve are consistent with the foreground pixels, the pixels are more likely to be foreground pixels. Thus the energy can be designed as the sum of consistency between the curve pixels and the inner region pixels:

$$E_{sf}^k = \int_{(x,y) \in \omega_{int}^k} f(I_k(x,y), g(\omega_{int}^k)) dx dy. \quad (4)$$

In above Eq. 4, $g(\omega_{int}^k)$ is for region metric, $f(I_k(x,y), g(\omega_{int}^k))$ is for measuring similarity between pixels and regions [8]. If the pixels are more similar to the inner region, the E_{sf}^k is larger, otherwise smaller.

The energy of background consistency of the k -th image: E_{sb}^k :

The background consistency of a single image is similar to foreground consistency of a single image, except it is used to ensure consistency of the background pixels. Because the background of images may be similar, we should take the energy of background consistency into account and λ_{sb} would be zero when the background of images are not similar. Thus, the energy of background consistency is designed as the consistency between the pixels inside the curve and the external region. The corresponding equation is:

$$E_{sb}^k = \int_{(x,y) \in \omega_{ext}^k} f(I_k(x,y), g(\omega_{ext}^k)) dx dy. \quad (5)$$

If the pixels are more similar to the outside region, the E_{sb}^k is larger, otherwise smaller.

The energy of foreground consistency of the k -th image and other images: E_f^k :

The foreground consistency of multiple images guarantees the consistency of the image foreground region with other images, which can be expressed as the sum of the image curve and the interior region of the rest of the images:

$$E_f^k = \sum_{j=1, j \neq i}^n \int_{(x,y) \in \omega_{int}^k} f(I_k(x,y), g(\omega_{int}^j)) dx dy. \quad (6)$$

If the curve of the image is consistent with most of the inner regions, the E_f^k is larger, and smaller in vice versa.

The energy of background consistency of the k -th image and the rest of the images: E_b^k :

Similar to foreground consistency of multi images, background consistency of multi-images can be expressed as the sum of the image curve and the outer region of the rest of images:

$$E_b^k = \sum_{j=1, j \neq i}^n \int_{(x,y) \in \omega_{ext}^k} f(I_k(x,y), g(\omega_{ext}^j)) dx dy. \quad (7)$$

If the external regions are more consistent with each other, the E_b^k is larger.

The saliency energy of the k -th image: E_{co}^k :

Saliency is used to obtain more foreground information. During this image saliency process, the pixels with greater saliency values are more likely to be foreground pixels, making the discrimination between foreground and background easier. Thus, the energy can be designed as the sum of the difference between the saliency value of pixel and the saliency value of region:

$$E_{co}^k = \int_{(x,y) \in \omega^k} h(J_k(x,y), l(\omega^k)) dx dy. \quad (8)$$

In above Eq. 8, $J_k(x,y)$ refers to the saliency value for each point of the k -th image, $l(\omega^k)$ represents the saliency value of measuring the region saliency, and $h(J_k(x,y), l(\omega^k))$ refers to pixel and region saliency differences. Therefore, if the saliency value of pixel is larger, then E_{co}^k is larger, and vice versa.

3 Optimization method

This section aims at model optimization and proposes an efficient level-set optimization method which is based on super-pixel, hierarchical segmentation and better termination.

3.1 Basic level-set optimization

The level-set optimization method of Osher and Sethian [11] can be used to minimize the energy function expressed in Eq. 2. Based on the level-set function $\varphi(x,y,t)$, Eq. 2 can be expressed as:

$$\begin{aligned} E(\varphi_k) = & \mu \int_{\Omega_k} H(\varphi_k) dx dy + \nu \int_{\Omega_k} \delta(\varphi_k) |\nabla \varphi_k| dx dy \\ & - \lambda_{sf} \int_{\Omega_k} f(I_k(x,y), g(\omega_{int}^k)) H(\varphi_k) dx dy \\ & - \lambda_{sb} \int_{\Omega_k} f(I_k(x,y), g(\omega_{ext}^k)) (1 - H(\varphi_k)) dx dy \\ & - \lambda_f \sum_{j=1, j \neq i}^n \int_{\Omega_k} f(I_k(x,y), g(\omega_{int}^j)) H(\varphi_k) dx dy \\ & - \lambda_b \sum_{j=1, j \neq i}^n \int_{\Omega_k} f(I_k(x,y), g(\omega_{ext}^j)) (1 - \\ & \quad H(\varphi_k)) dx dy \\ & - \lambda_{co} \int_{\Omega_k} h(J_k(x,y), l(\omega_{int}^k)) H(\varphi_k) dx dy. \quad (9) \end{aligned}$$

Here, Ω_k is the k -th image in the image domain, φ_k is the k -th image level-set function, $\delta(x)$ is the Dirac function, $H(x)$ is the Heaviside function. Subsequently, $\delta(\varphi_k)$ stands for C_k , $H(\varphi_k)$ is for the inner region and $1 - H(\varphi_k)$ is for the outer region. The Dirac and Heaviside functions are defined as follows:

$$\begin{cases} H(x) = \frac{1}{2} \cdot (1 + \frac{2}{\pi} \arctan \frac{x}{\varepsilon}) \\ \delta(x) = \frac{dH(x)}{dx} = \frac{1}{\pi} \cdot \frac{\varepsilon}{\varepsilon^2 + x^2} \end{cases}. \quad (10)$$

Energy function expressed in Eq. 9 can be optimized using Euler Lagrange by keeping $f(I_k(x,y), g(\omega_{ext}^k))$ and φ_k as independent and keeping $g(\omega_{int}^k)$ and $g(\omega_{ext}^k)$ as constant.

Set $\Delta t = 1$, we can get the iterative formula of the level-set function:

$$\begin{aligned} \varphi_k^{N+1} = & \varphi_k^N + \delta(\varphi_k) (-\mu + \nu \cdot \text{div} \left(\frac{\nabla \varphi_k}{|\nabla \varphi_k|} \right) \\ & + \lambda_{sf} f(I_k(x,y), g(\omega_{int}^k)) \\ & - \lambda_{sb} f(I_k(x,y), g(\omega_{ext}^k)) \\ & + \lambda_f \sum_{j=1, j \neq i}^n f(I_k(x,y), g(\omega_{int}^j)) \\ & - \lambda_b \sum_{j=1, j \neq i}^n f(I_k(x,y), g(\omega_{ext}^j)) \\ & - \lambda_{co} h(J_k(x,y), l(\omega_{int}^k)). \quad (11) \end{aligned}$$

Finally, co-segmentation algorithm based on the active contour using image saliency is obtained and the iterative equation of the level-set function is expressed in Eq. 11.

Given a group of images $I = \{I_1, I_2, \dots, I_n\}, n \geq 2$, and initial contour sets $C^0 = \{C_1, C_2, \dots, C_n\}$, we can get the initial level-set $\varphi^0 = \{\varphi_1^0, \varphi_2^0, \dots, \varphi_n^0\}$. At each step of the iteration, we calculate the energy-value of each item, then update φ_k^{N+1} values according to Eq. 11, and finally, calculate internal area $g(\omega_{int}^k)$ and external area $g(\omega_{ext}^k)$. If the convergence conditions are obtained, algorithm is terminated. When working with level-sets and Dirac delta functions, a standard procedure is to reinitialize to the signed distance function to its zero-level curve. This prevents the level-set function to become too flat, or it can be seen as a rescaling and regularization [10].

3.2 An efficient level-set optimization

If the algorithm derived above is directly calculated on pixels, the calculation will be very expensive. In order to speed up the optimization, the following strategies, including super-pixels, hierarchical computation and convergence judgment, are introduced.

3.2.1 Super-pixels

To accelerate the computation, super-pixels based computation is performed to reduce the computational burden drastically. Taking 200×200 image as an example, the calculation directly on original pixels computes the whole 40000 pixels totally, contrarily number reduced to 1000, or 1/20th of the original level, if super-pixels are used. In this paper, we use Simple Linear Iterative Clustering (SLIC) method [12] to generate super-pixels since this method is commonly used and has good results, and the default super-pixels number is 1000.

3.2.2 Hierarchical Computation

From the view of down sampling, hierarchical data processing can be used to speed up computation. Hierarchical computation includes several layers and processes layer by layer, each layer processes one size of the original image and runs co-segmentation algorithm until convergence. The hierarchical structure starts from small size of images with the quick convergence to the object boundary and then propagates the segmentation from the images with small size to the images with large size. As initial contour of the last layer is somehow close to the final results, it is possible to achieve faster convergence. This whole reducing scheme is called cascade acceleration and it significantly increases the speed of the algorithm.

3.2.3 Convergence Judgment

In general, the algorithm sets an iteration number, and continues until the number of iterations is achieved. Large iteration number guarantees the convergence but on the huge computational cost while small iteration number, for avoiding the computational cost, hinges on the uncertainty of satisfactory convergence. In this paper, to our algorithm judges whether it is converged during iterative process and terminate timely. Specifically, the algorithm is converged if φ_k^{t+1} tends to remain stable and in this way we can set a relatively large number of iterations without worrying of computational cost. In this way, we can accelerate the algorithm and ensure good segmentation results at the same time.

To sum up, our algorithm of co-segmentation is summarized in Algorithm 1.

4 Experiments

In order to evaluate the proposed co-segmentation algorithm, we conducted the experiments on commonly used datasets and compared our algorithm with other competitive methods.

4.1 Experimental Setup

4.1.1 Image dataset

The iCoseg image dataset [13] and MSRC image dataset [14] are used in our experiments, which are often used in co-segmentation field. The iCoseg image dataset is established by Batra et al. [5], which contains a total of 37 categories. Each category has several to dozens of images. The MSRC image dataset was established by Criminisi et al. [14]. This image dataset includes 23 categories, and each category has dozens of images.

The effectiveness of results obtained by our method is evaluated with ground truths. The error rates are measured, which is defined as the number of misclassified pixels divided by the number of the total pixels.

4.1.2 Parameter setting

Algorithm 1. Our co-segmentation algorithm based on active contour and image saliency (ACIS)

-
- Step 1. Set initial parameters μ , ν , λ_{sf} , λ_{sb} , λ_f , λ_b and λ_{co} .
 - Step 2. Extract the information of each image I, super-pixel information and corresponding saliency information J.
 - Step 3. Initialize the contour of each image $C^0 = \{C_1, C_2, \dots, C_n\}$.
 - Step 4. The original image, the corresponding super-pixel information and saliency information are set to different sizes.
 - Step 5. Using super-pixels to achieve co-segmentation of different image sizes and enter Step 6 until all sizes are processed.
 - Step 5.1. Given C^0 , the initial level set function φ^0 is obtained according to [10].
 - Step 5.2. According to φ^0 , get internal area representation $g(\omega_{int}^k)$ and external regional representation $g(\omega_{ext}^k)$.
 - Step 5.3. Iterative calculation.
 - Step 5.3.1. Calculate each energy item according to Eq. 11.
 - Step 5.3.2. The level set function of each image φ_k^{t+1} is updated according to φ_k^t and Eq. 11.
 - Step 5.3.3. Reinitialize to the signed distance function to update φ_k^{t+1} .
 - Step 5.3.4. Based on φ_k^{t+1} , calculate internal area $g(\omega_{int}^k)$ and external area $g(\omega_{ext}^k)$.
 - Step 5.3.5. Judging convergence condition, if meet the convergence condition enter Step 5, otherwise enter Step 5.3.1.
 - Step 6. Output segmentation results.
-

The parameters of our algorithm are set up by serious experiments. The actual values leading to the reported results are listed as follows.

- For the initial contour, $\alpha = 0.2 \cdot \min\{\text{height}, \text{width}\}$, height stands for the height of an image, width stands for the width of an image.
- In Eq. 11, considering the relationship between each item, we here set $\mu = 0.01$, $\nu = 0.05$, $\lambda_{sf} = \lambda_f \in [0.5, 1.1]$, $\lambda_{sb} = \lambda_b = 1$, $\lambda_{co} \in [0.1, 0.3]$.
- The parameter for measuring whether pixels are similar is $\beta = 15$ [8].
- In the experiments, the level number of the cascade acceleration is three, the expansion scale is $\{0.1, 0.5, 1\}$ and the corresponding iteration numbers are $\{100, 30, 30\}$.

4.1.3 Compared Methods

GrabCut [15] is a very famous image segmentation method, we compared this method to show ACIS has better performance. We also compared our method with other co-segmentation methods including CoSB [16] and CoSand [17]. CoSB is based on active contour and CoSand is based on thermal expansion. We set default

parameter values according to the original papers [15], [16], [17].

4.2 Evaluation of saliency methods

Saliency item in Eq. 11 greatly affects the segmentation results. As the weight of saliency item is increased, the segmentation results become more close to saliency results. In this paper, we consider four saliency methods, including GS [18], MR [19], SF [20] and wCtr [21]. We compared them through careful evaluation. The results are shown in Fig. 1. The first row in Fig. 1 represents the original images and the second row represents the ground truth while the remaining rows represent the saliency images extracted by GS [18], MR [19], SF [20], and wCtr [21], respectively.



Figure 1. First row is the original image, second row is the ground truth, and other rows are the saliency extraction results using GS, MR, SF and wCtr, respectively, on iCoseg image dataset.

From the results shown in Fig. 1, we can see that extraction results of the GS method for the foreground saliency are good, but some unnecessary backgrounds also have high saliency value. MR method obtains lower value and cannot accurately extract the foreground saliency. The SF method fails to accurately extract the foreground saliency in the bear class. Meanwhile, the extraction results of the wCtr method are more accurate and robust than other methods. Table 1 shows the comparison of the different saliency detection methods on the error rate with different thresholds. We can see that in each threshold and in each category, wCtr has the lowest error rate. Therefore, we chose the wCtr method for the extraction of the saliency in this paper.

Table 1. The error rates comparison of different saliency detection methods.

Thresholds	Category	GS [18]	MR [19]	SF [20]	wCtr [21]
Mean value	Goose	0.159	0.227	0.054	0.050
	Bear	0.204	0.242	0.312	0.152
1.2*Mean value	Goose	0.120	0.200	0.054	0.047
	Bear	0.204	0.233	0.316	0.159
Fixed value(127)	Goose	0.065	0.117	0.096	0.041
	Bear	0.198	0.192	0.316	0.179
Avg		0.158	0.202	0.191	0.105

4.3 Results on iCoseg Image dataset

The first part is the subjective comparison of co-segmentation results. Fig. 2 shows the comparison of different segmentation methods on the iCoseg image dataset. From left to right and from top to bottom, the classes are Sox players, Pyramids, Brighton and Helicopter, respectively.



Figure 2. Comparison of different segmentation methods on iCoseg image dataset.

For each class, the first row contains the original images and the second row contains ground truth images. The third row shows segmented images using GrabCut [15]. Results shown in the fourth and fifth rows are segmented results using CoSB and CoSand, respectively. Results in the sixth row are obtained using our method.

According to Fig. 2, for the class of Sox Players, the method CoSand accurately segments the foreground image from the grass but in due process over-segment the foreground objects, resulting in loss of information. In the contrary, GrabCut and CoSB have poor results. Since foregrounds and backgrounds of Pyramids class are very similar, Grab Cut and CoSB failed to distinguish Pyramid from desert. Somehow, our method can separate them in some degree. Moreover, for the Brighton class, the results of segmentation obtained by CoSB are good, but GrabCut has bad results and results of CoSand appear over-segmented. For the Helicopter class, GrabCut and CoSB failed on some images and CoSand also appears over-segmented. In general, the effects and robustness of our method, ACIS, are more excellent compared with other methods. Table 2 shows the comparison of the different methods and our method on the error rate.

GrabCut is based on single image segmentation. So it performs well in skating and gymnastics class because the background is relatively simple, but its performance

deteriorates on other classes. CoSB performs well in Brighton class. Because this method only considers the consistency of foreground and background without adding more foreground information. So in many cases of co-segmentation, especially under the complex background, it is difficult to define a suitable energy function. The CoSand, a thermal diffusion method, has good results on the classes of skating and hot balloons, but due to heavily relying on the color, the method will fail easily if the foreground has different colors. ACIS is based on active contour using saliency, active contour guarantees the curve close to the edge of the object and saliency ensures stronger robustness. Therefore, our method has good co-segmentation results in most classes.

Table 2. The comparison of error rates on iCoseg image dataset.

Category	GrabCut [15]	CoSB [16]	CoSand [17]	ACIS (Our)
Bear	0.241	0.331	0.105	0.073
Sox players	0.283	0.449	0.027	0.023
Stonehenge	0.239	0.230	0.075	0.064
Pyramids	0.246	0.380	0.266	0.130
Elephants	0.131	0.285	0.120	0.121
Helicopter	0.142	0.054	0.071	0.011
Cheetah	0.094	0.092	0.294	0.065
Brighton	0.104	0.016	0.047	0.016
Margate	0.218	0.111	0.264	0.086
Gymnastics	0.013	0.053	0.038	0.020
Skating	0.019	0.184	0.019	0.028
Balloons	0.050	0.054	0.028	0.037
Avg	0.148	0.187	0.113	0.056

4.4 Results on MSRC Image dataset

Fig. 3 shows the comparison of different segmentation methods on MSRC image dataset. From left to right and from top to bottom, the classes are cow, face, sign and cat, respectively.

For the cow class, GrabCut and ACIS performed well despite of different foreground colors, while CoSB and CoSand performs badly because of the different colors of foregrounds. For the face class, ACIS has the best segmentation effects, while other methods do not have satisfactory results due to many interferences and complexity of background. For the sign class, all the methods have relatively good segmentation results, but CoSand takes the sign post as the foreground in the 2nd image while the sign post actually belongs to the background. For the cat class, ACIS also has the best segmentation results. Comparison of the error rates of different segmentation methods are given in the following Table 3.

As shown in Table 3, ACIS has the best average value. GrabCut has relatively lower error rate in most classes while CoSB and CoSand did not perform well on the dataset. The reason is that the foreground and background are relatively complex and different colors in foreground

make it more difficult. ACIS has the lowest error rate in the classes of the face and cat and the average error rate is lower. Thus, it is proved from aforementioned results that our method is effective.

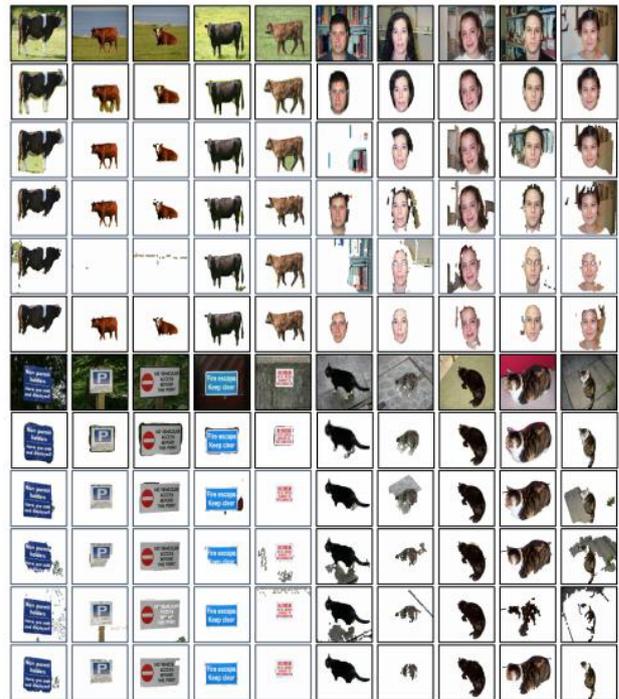


Figure 3. Comparison of different segmentation methods on MSRC image dataset.

Table 3. The comparison of error rates on MSRC image dataset.

Category	GrabCut [15]	CoSB [16]	CoSand [17]	ACIS (Our)
Cow	0.063	0.070	0.122	0.068
Face	0.232	0.175	0.278	0.126
Bicycle	0.204	0.235	0.316	0.239
Sign	0.038	0.082	0.077	0.042
Bird	0.039	0.151	0.137	0.072
Chair	0.111	0.198	0.281	0.116
Cat	0.091	0.125	0.152	0.055
Avg	0.111	0.148	0.195	0.103

The main difference between ACIS and CoSB is that CoSB is based on active contour while ACIS is based on active contour and image saliency. So the above experiments show that introducing saliency information can improve co-segmentation results. In conclusion, compared with other methods, our method yielded better results and demonstrated the significance of using image saliency in active contour.

4.5 Efficiency

Except the comparison of error rates, ACIS is also an efficient image co-segmentation algorithm because we adopt an efficient level-set optimization method. Table 4

shows the comparisons of computation cost on MSRC image dataset. From the results we can see that our optimized level-set is much faster than the original level-set.

Table 4. Computation cost comparison of level-set and optimized level-set on MSRC image dataset.

Category	Level-set [10]	Optimized level-set
Cow	>1000s	115.2s
Face	>1000s	91.6s
Bicycle	>1000s	298.6s
Sign	>1000s	87.8s
Bird	>1000s	37.0s
Chair	>1000s	37.8s
Cat	>1000s	32.4s
Avg	>1000s	100.0s

5 Conclusion

This paper has proposed an efficient co-segmentation algorithm based on active contour and image saliency. Firstly, we construct a new image co-segmentation model by introducing the saliency information into the active contour framework. Secondly, we proposed an efficient algorithm which is using level-set method with super-pixels, hierarchical segmentation, and better termination to optimize the model above. In order to let the algorithm has more robust and better segmentation effect, the choice of saliency method is discussed using rigorous evaluation. Finally, we conducted a series of experiments on iCoseg and MSRC datasets. Compared with other methods, our method yielded better co-segmentation results and computationally inexpensive. In future work, we will adopt more accurate foreground information, such as co-saliency, since saliency is just for a single image.

References

1. C. Rothe, V. Kolmogorov, T. Minka, and A. Blake, "Cosegmentation of Image Pairs by Histogram Matching - Incorporating a Global Constraint into MRFs," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **1**, 993-1000 (2006).
2. L. Mukherjee, V. Singh, R. And Dyer C, "Half-Integrality based Algorithms for Cosegmentation of Images," *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2028-2035 (2009).
3. S. Hochbaum D, and V. Singh. "An efficient algorithm for Co-segmentation," *IEEE International Conference on Computer Vision*, 269-276 (2009).
4. S. Vicente, V. Kolmogorov, and C. Rother, "Cosegmentation revisited: Models and optimization," *ECCV*, 465-479 (2010).
5. A. Joulin, F. Bach, and J. Ponce, "Discriminative clustering for image co-segmentation," *IEEE Conference on Computer Vision & Pattern Recognition* *IEEE Computer Society*, 1943-1950 (2010).
6. D. Batra, A. Kowdle, and D. Parikh D, "iCoseg: Interactive co-segmentation with intelligent scribble guidance," *Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3169-3176 (2010).
7. S. Vicente, C. Rother, and V. Kolmogorov, "Object cosegmentation," *IEEE Conference on Computer Vision & Pattern Recognition*, **42**, 2217-2224 (2011).
8. F. Meng, H. Li, and G. Liu, "Image co-segmentation via active contours," *IEEE International Symposium on Circuits & Systems*, 2773-2776 (2012).
9. F. Meng, H. Li, and G. Liu, "Image Cosegmentation by Incorporating Color Reward Strategy and Active Contour Model," *Cybernetics IEEE Transactions on*, **43**, 725-737 (2013).
10. F. Chan T and A. Vese L, "Active contours without edges," *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, **10**, 266-277 (2001).
11. S. Osher and A. Sethian J, "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, **79**, 12-49 (1988).
12. R. Achanta, A. Shaji, K. Smith, et al. "SLIC Superpixels Compared to State-of-the-art Superpixel Methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **34**, 2274-2282 (2012).
13. Available at : http://chenlab.ece.cornell.edu/projects/touch-coseg/CMU_Cornell_iCoseg_dataset.zip, last accessed: 04/11/2015.
14. Available at : http://research.microsoft.com/en-us/um/people/antcristm/data_objrec/msrc_objcategimagedatabase_v2.zip, last accessed: 04/11/2015.
15. C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: Interactive Foreground Extraction Using Iterated Graph Cuts," *ACM Transactions on Graphics*, **23**, 309-314 (2004).
16. F. Meng, H. Li, K. Ngan, et al. "Cosegmentation from similar backgrounds," *IEEE International Symposium on Circuits & Systems*, 353 - 356 (2014).
17. G. Kim, P. Xing E, L. Fei-Fei, et al. "Distributed cosegmentation via submodular optimization on anisotropic diffusion," *Computer Vision (ICCV), 2011 IEEE International Conference on*, 169-176 (2011).
18. Y. Wei, F. Wen F, W. Zhu, et al. "Geodesic Saliency Using Background Priors," *Lecture Notes in Computer Science*, **7574**, 29-42 (2012).
19. C. Yang C, L. Zhang, H. Lu, et al. "Saliency Detection via Graph-Based Manifold Ranking," *IEEE Conference on Computer Vision & Pattern Recognition*, **9**, 3166-3173 (2013).
20. A. Hornung, Y. Pritch, P. Krahenbuhl, et al. "Saliency filters: Contrast based filtering for salient

- region detection," *Cvpr. IEEE Computer Society*, 733-740 (2012).
21. W. Zhu, S. Liang, Y. Wei, et al. "Saliency Optimization from Robust Background Detection," *IEEE Conference on Computer Vision & Pattern Recognition*, 2814 – 2821 (2014).