

# Environment-Sensitive cloning in images

Yun Zhang · Ruofeng Tong

© Springer-Verlag 2011

**Abstract** We present an Environment-Sensitive image cloning technique which improves the previous gradient-based methods by taking into account the content of target scene. We create a reference image to represent the global feature of target image which could be further diffused into the cloned patch, and modify the diffusion process to ensure that the cloning result is seamless and natural. Specifically, we figure out an efficient solution based on Mean-Value Coordinates (MVC) to deal with the hybrid boundary, and construct a general model based on MVC to implement our image cloning, which is further applied to video cloning. Experimental results demonstrate the effectiveness of our Environment-Sensitive cloning.

**Keywords** Environment-Sensitive · Global feature · Mean-Value Coordinates (MVC)

## 1 Introduction

Image cloning has always been a hot topic in image editing, which can be described as creating a new image by cloning a source image patch into the target image seamlessly and naturally. The common method to perform image cloning is a simple process of cut-and-paste with interactive image editing tools (e.g. PhotoShop). However, it is labor-intensive and difficult for users to select the source image patch exactly. Moreover, it is hard to modify appearance of

the cloned patch to match the target scene. To address this problem, the gradient-based method is proposed. Compared with the simple cut-and-paste, the advantages of this method are two-fold. Firstly, the selection of source image patch is easy without too much skills and labors. Secondly, the intensities of cloning region can be reconstructed to be consistent with the target scene by solving a partial differential equation (PDE).

In general, the gradient-based method is effective in accomplishing seamless image cloning when textures are similar in the source and target image near the cloning boundary. However, this method only considers the local feature which refers to the cloning boundary of the target image for seamless cloning while ignoring the global feature of the scene. As shown in Fig. 3(d), the bear is seamlessly cloned to the beach under the sunset using the gradient domain method. But when judged globally, it appears to be inconsistent with the target scene. This is owing to the fact that the gradient domain method only diffuses the local feature into the entire cloned patch without considering the global feature. We believe that the global feature is an important factor for high-quality image cloning.

In this paper, we propose an Environment-Sensitive cloning method which pays attention to global feature of the target scene. The main idea of our method is to construct a reference image which can represent the global feature of the target scene and influence the appearance of cloned patch. The local and global features are diffused into the cloning area under the control of alpha matte. This ensures that the cloned patch is seamless and consistent with the target scene when pasting. In addition, we conceive an efficient method to perform image composition with the hybrid boundary [3], and construct a general model to implement our image cloning. At last, we apply our environment-

---

Y. Zhang · R. Tong (✉)  
Department of Computer Science and Technology, Zhejiang  
University, Hangzhou 310027, China  
e-mail: [trf@zju.edu.cn](mailto:trf@zju.edu.cn)

Y. Zhang  
e-mail: [zhangyun\\_zju@zju.edu.cn](mailto:zhangyun_zju@zju.edu.cn)

sensitive image cloning to video by introducing the up-to-date Video cutout [2] and matting technique [14].

The rest of the paper is organized as follows. After briefly summarizing the related work in Sect. 2, we discuss our Environment-Sensitive image cloning, and then apply it to video cloning in Sect. 3. Experimental results and discussions are given in Sects. 4 and 5, and conclusions in Sect. 6.

## 2 Related work

*Gradient domain approaches* Recently, the gradient domain approaches are very popular to perform seamless cloning, which make use of the psychological discovery [13] that the human visual system (HVS) is much more sensitive to local contrasts than to slow changes in luminance and chrominance. Poisson image editing [16] has been proposed as an effective approach for seamless image composition, which constructs a harmonic membrane to smoothly interpolate the discrepancies between the boundary of source patch and target image to the entire cloned area. However, this method involves solving a large sparse linear system which is time and memory consuming. Wang et al. [19] proposed a seamless video composition technique by 3D integration in the gradient domain, which is a natural extension of Poisson image editing. Jia et al. [10] further improved the Poisson blending by calculating an optimized boundary to produce better results. Farbman et al. [8] proposed to take Mean-Value Coordinates (MVC) rather than solving a large linear system to perform image composition, and the performance is advantageous in terms of speed, ease of implementation, memory footprint, and parallelization. Chen et al. [3] proposed a novel hybrid image cloning method, which combines improved Poisson and alpha blending with optimized boundary. Very recently, the idea of combining the gradient-based cloning and alpha matte has been proposed to eliminate smudging and discoloration artifacts [10, 12, 22]. Although the cloning methods above are effective to produce seamless cloning, the results are independent of the global illumination condition of the scene and the content of the cloned patch. Therefore, the intensity changes in the cloned area may be unnatural. With this in mind, Ding et al. [7] proposed the content-aware cloning, in which the diffusion process can be easily controlled by matting technique. This helps to produce seamless cloning while preserving the color appearance of the important part in the selected region. However, it still fails to cope with illumination inconsistency between the source patch and the target scene. Unlike previous cloning methods, we take into account both local and global features of the target scene in the cloning process to yield more realistic and visual appealing results.

*Image and video matting* In our image and video cloning, matte plays an important role. The boundary of the cloned patch is given by the *trimap* used in alpha matting. And the soft alpha matte is used to control the diffusion process in the cloning area. Since the matte used in gradient-based method does not determine the image intensities directly, it is unnecessary to extract an accurate one. Recently, tremendous progress has been achieved on interactive matting of images, which is summarized in [21]. In current image matting systems [5, 9, 14, 17, 18, 20], users are required to specify scribbles or *trimap* indicating foreground, background, or unknown pixels, and the matte of unknown pixels is estimated based on the known information. In our image cloning system, the closed-form matting technique [14] is used to generate high-quality matte which is derived by solving a sparse linear system. For objects with complex topologies, e.g. uncollected regions, users can choose the color based tool such as affinity-based edit propagation [15, 23] and boundary band based RepFinder [4] to complete this task. When 2D matting is extended to 3D video frames, the problem becomes more challenging. Armstrong et al. [1] extends 2D graph cut directly to 3D video frames. Chuang et al. [6] applied Optical flow to propagate the segmentation between frames. The recently proposed Video SnapCut [2] employs a set of local classifiers to extract the foreground from complex scenes. We further improve the efficiency and user interaction of Video SnapCut system in our application of video composition.

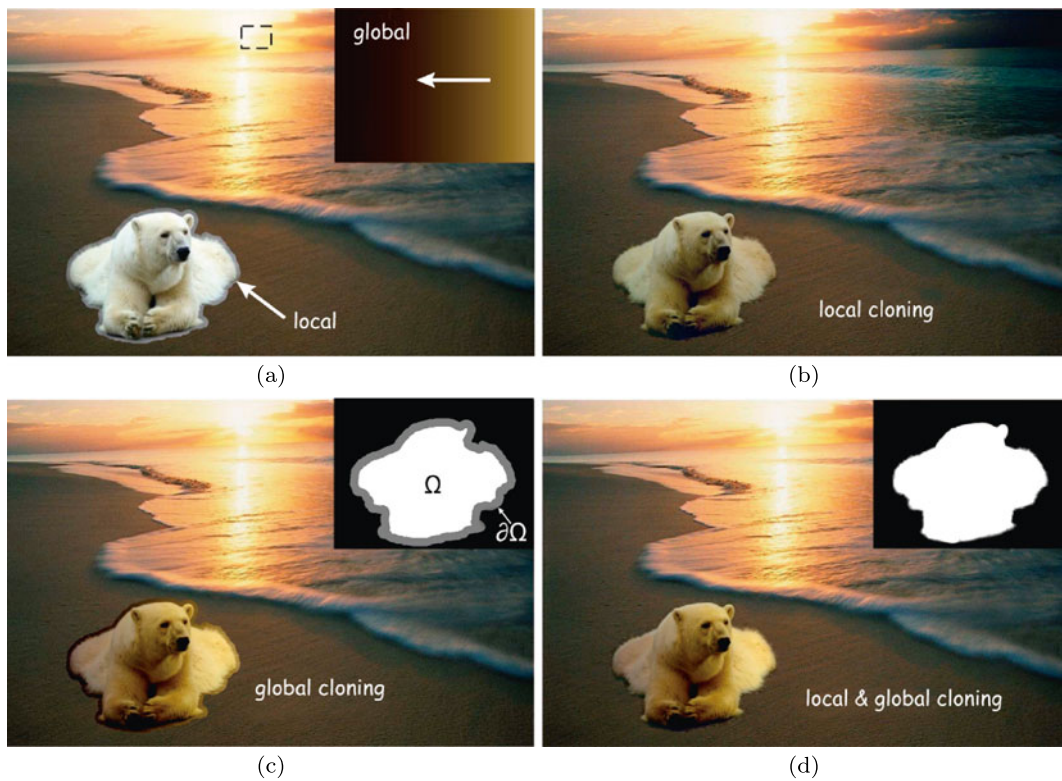
## 3 Environment-Sensitive image cloning

In our approach, we first create a reference image to represent the global feature of the target scene. Then, both the global and local features of the target image are diffused into the cloning area under the control of alpha matte, which ensures that the final result is natural and consistent with the target scene.

We give the notations of our image cloning as follows: Let  $S \subset \mathbb{R}^2$ ,  $T \subset \mathbb{R}^2$  and  $R \subset \mathbb{R}^2$  be the domains of source, target and reference images. The intensities of these images are denoted by  $g : S \rightarrow \mathbb{R}$ ,  $f^* : T \rightarrow \mathbb{R}$  and  $\xi : R \rightarrow \mathbb{R}$ , respectively. Let  $\Omega \subset \mathbb{R}^2$  be the region to be cloned with the boundary  $\partial\Omega$ .

### 3.1 Global feature extraction

In general, the appearance of an object in the scene is mainly determined by the illumination conditions. However, regarding a single image, there are no simple methods to extract and represent the illumination features [11]. When we observe a single image carefully, a lot of cues in it can tell us its illumination features (e.g. the position of the sun can be



**Fig. 1** An example of Environment-Sensitive image cloning: **(a)** Source and target image (global and local features are indicated in the target scene); **(b)** result by diffusing the local feature into the cloning area; **(c)** result by diffusing the global feature into the cloning area

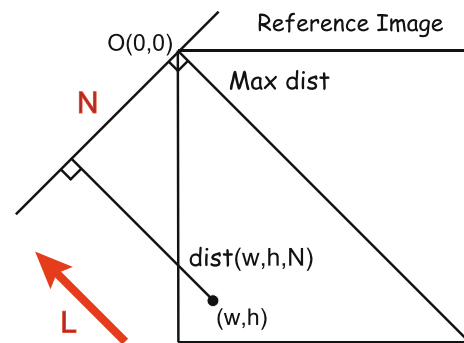
**(trimap** which is used to specify the cloning boundary is in the *top-right corner*); **(d)** result by our method which diffuses the global and local features of the scene (Alpha matte which is used to control the diffusion is in the *top-right corner*)

indicated by the shadows on the ground or the shading on vertical objects, and the appearance of the existing objects in the scene can also tell us the illumination conditions). With this in mind, we make use of these cues in a single image to extract the illumination features of the target scene by human interactions, and create a reference image  $\xi$  to represent the extracted features, which can then be easily diffused into the cloned patch by gradient-based method. As shown in Fig. 1(a), the light direction is marked by a white arrow, which indicates that the right part of  $\xi$  should be brighter, while the left part is darker. The size of  $\xi$  is the same as the boundary box of the cloned patch.

We focus on constructing the reference image  $\xi$  when there is a prevailing illumination from a certain direction in the target scene. The construction of  $\xi$  is in *LAB* color space which separates the lightness *L* and color component *A*&*B* effectively. The lightness *L* can be set as shown in Fig. 2, after simple interactions, the lightness distribution of the reference image can be expressed by

$$\xi_L(w, h) = \left( \frac{\text{dis}(w, h, N)}{d_M} \right)^\tau * V_M \quad (1)$$

where *N* is a line perpendicular to the direction of red arrow, and the position of *N* is indicated in Fig. 2.  $V_M$  is



**Fig. 2** Distance definition for creating a reference image

the maximum lightness, and  $\text{dis}(w, h, N)$  is the distance of pixel  $(w, h)$  to *N*, while  $d_M$  is the distance from the farthest pixel of  $\xi$  to *N*.  $\tau$  is used to control the illumination changes in  $\xi$ . See Fig. 2 for the distance definition in details.

Next, we consider the color style of the reference image. We first transform the target image from *RGB* to *LAB* color space, and then select a region in the target image which represent the color style of the target scene. As shown in Fig. 1(a), we select a region in the sunset with a rectangle, and then compute the mean values of color component

$Mean_A$  and  $Mean_B$  in it. For all pixels in the reference image, their  $A$  &  $B$  components are set to  $Mean_A$  and  $Mean_B$ , respectively.

After calculating the lightness  $L$  and color components  $A$  &  $B$ , the reference image  $\xi$  is constructed in  $LAB$  which is then converted to the  $RGB$  color space. The top-right corner of Fig. 1(a) is an example of reference image constructed by our method, which characterizes the global feature of the target scene.

### 3.2 Approach of Environment-Sensitive image cloning

Our cloning approach combines the local and global features of the target scene. As shown in Fig. 1(a), the global feature is represented by a reference image  $\xi$ , while the local feature refers to the pixels' color on the cloning boundary of target image.

Since the essence of gradient-based cloning is to construct a harmonic interpolant  $R(x)$  which diffuses the discrepancies on the boundary of the source and target image into the cloned patch, the final result can be simply expressed as

$$f(x) = g(x) + R(x). \tag{2}$$

When the global feature is diffused into the cloning area by constructing a harmonic interpolant  $r_1(x)$ , as shown in Fig. 1(c), the cloned patch appears to be consistent with the target scene. However, the visible seam exists near the cloning boundary. To remove the seam, the local feature should be taken into account by constructing a harmonic interpolant  $r_2(x)$  to diffuse the discrepancies on the cloning boundary of target and source images. See Fig. 1(b), the cloned patch is pasted into the target scene seamlessly while ignoring the global illumination feature.

The unsatisfactory results above arise because the global and local features are considered separately. We believe that these features should collaborate to yield satisfactory results. Thus, we construct a new interpolant  $R(x)$  which is a linear combination of the interpolants used in global and local diffusion. For important region of the cloned patch whose appearance should be changed to match the target scene, a larger factor should be given to the global interpolant. While for region near the cloning boundary, the visible seam should be smoothly diffused, so that the local interpolant is given a larger factor. Thus, it is natural to express the factor by alpha matte  $\alpha : S \rightarrow [0, 1]$ .  $\alpha$  represents the extent to which a pixel's color is influenced by the local and global features of the scene. Thus the interpolant  $R(x)$  can be expressed by

$$R(x) = \alpha(x)r_1(x) + (1 - \alpha(x))r_2(x) \tag{3}$$

where  $r_1(x)$  and  $r_2(x)$  are the interpolants of global and local diffusion, respectively. As shown in Fig. 1(d), when the

interpolants are constructed by (3), the source patch can be pasted into the target scene seamlessly and naturally.

### 3.3 Implementation of Environment-Sensitive image cloning

Unlike the previous cloning approaches, we do not have to specify the region to be cloned by drawing a boundary. Since we have to create a *trimap* for calculating alpha matte of the source image, the cloning boundary  $\partial\Omega$  can be given by the outer boundary of *trimap*, as shown in Fig. 1(c).

As the traditional gradient cloning involves solving a PDE, which is time and memory consuming, we prefer to employ Mean-Value Coordinates (MVC) [8] to approximate a harmonic membrane without solving a large linear system.

Given the cloning boundary (the points are ordered counter clockwise)  $\partial\Omega = (P_0, P_1, \dots, P_{n-1}, P_0)$ ,  $P_i \in \mathbb{R}^2$ . The MVC of a point  $x \in \mathbb{R}^2$  inside the region  $\Omega$  are given by

$$\lambda_i(x) = \frac{w_i}{\sum_{i=0}^{n-1} w_i} \quad i = 0, \dots, n - 1 \tag{4}$$

where

$$w_i = \frac{\tan(\alpha_{i-1}/2) + \tan(\alpha_i/2)}{\|P_i - x\|} \tag{5}$$

$\alpha_i$  is the angle  $\angle P_i x P_{i+1}$  (see details in [8]).

After calculating the MVC of the points in  $\Omega$ , the discrepancies on the boundary of source and target image can be smoothly diffused into the entire region  $\Omega$  by constructing a harmonic interpolant [8]. Guided by MVC, the harmonic interpolants  $r_1(x)$ ,  $r_2(x)$  can be expressed as

$$r_1(x) = \sum_{i=0}^{n-1} \lambda_i(x)(\xi(P_i) - g(P_i)), \tag{6}$$

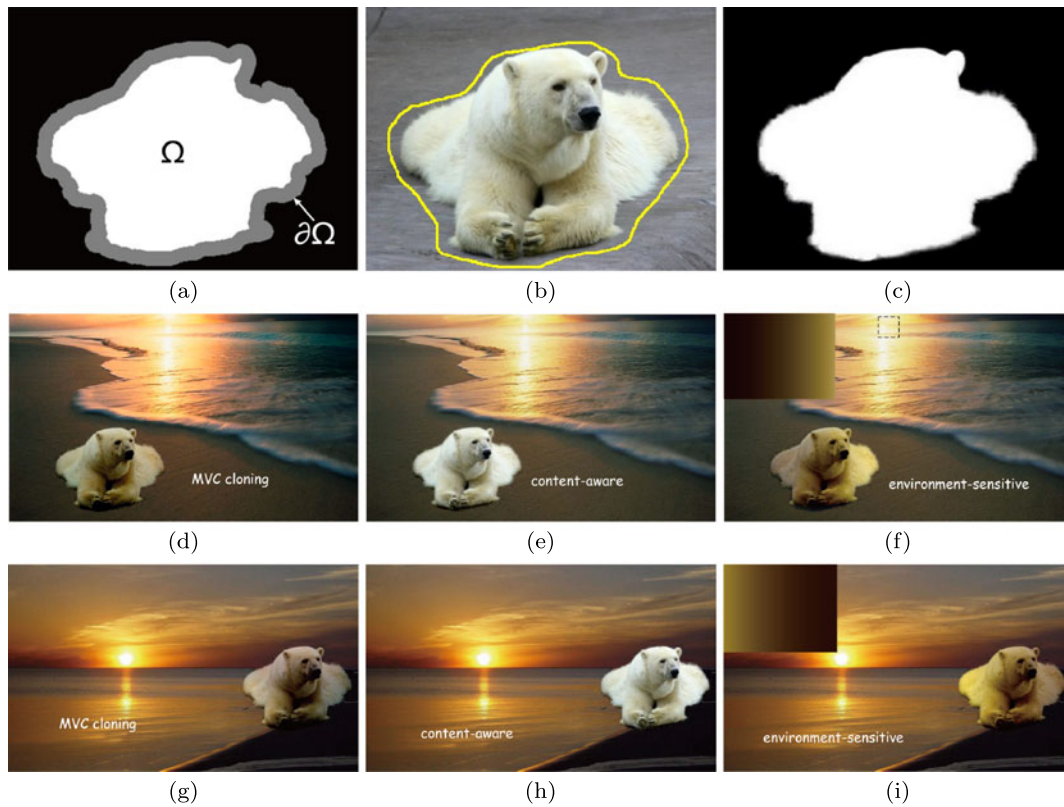
$$r_2(x) = \sum_{i=0}^{n-1} \lambda_i(x)(f^*(P_i) - g(P_i)). \tag{7}$$

As the cloning boundary does not change in the diffusion process of global and local features, the mean-value coordinates  $\lambda_i(x)$  can be shared by  $r_1(x)$  and  $r_2(x)$ , which takes  $\partial\Omega = (P_0, P_1, \dots, P_{n-1}, P_0)$ ,  $P_i \in \mathbb{R}^2$  as the boundary.

After MVC has been introduced into (2) and (3), the final cloning result can be expressed by

$$f(x) = g(x) + \sum_{i=0}^{n-1} \lambda_i(x)(\alpha(x)\xi(P_i) + (1 - \alpha(x))f^*(P_i) - g(P_i)). \tag{8}$$

Equation (8) is not only an implementation of our Environment-Sensitive cloning, but a general model for image cloning. Note that the color discrepancies on the cloning



**Fig. 3** Pasting a bear into the scenes with dominant illumination conditions: **(a)** The *trimap* of source image; **(b)** the source patch specified by the *trimap*; **(c)** alpha matte of source image patch; **(d)–(f)** composite results using direct MVC [8], content-aware [7], and Environment-

Sensitive cloning methods (the target scene is featured with sunset from *the right side*); **(g)–(i)** experiments similar to **(d)–(f)**, the target scene here is characterized by the sunset from *the left side*

boundary is expressed by  $(\alpha\xi + (1 - \alpha)f^* - g)$ , which means that the target boundary condition is determined by both the reference image  $\xi$  and target image  $f^*$ .  $\alpha$  is used to control the target boundary condition in the following way: the interpolants of pixels with higher  $\alpha$  value will be mainly determined by the reference image  $\xi$ . While for pixels with lower  $\alpha$  value, their interpolants are influenced by the target image  $f^*$ . Thus, the global feature will be diffused into the important part of the cloned patch, while the unimportant part (e.g. the region near the cloning boundary) will blend with the target image seamlessly.

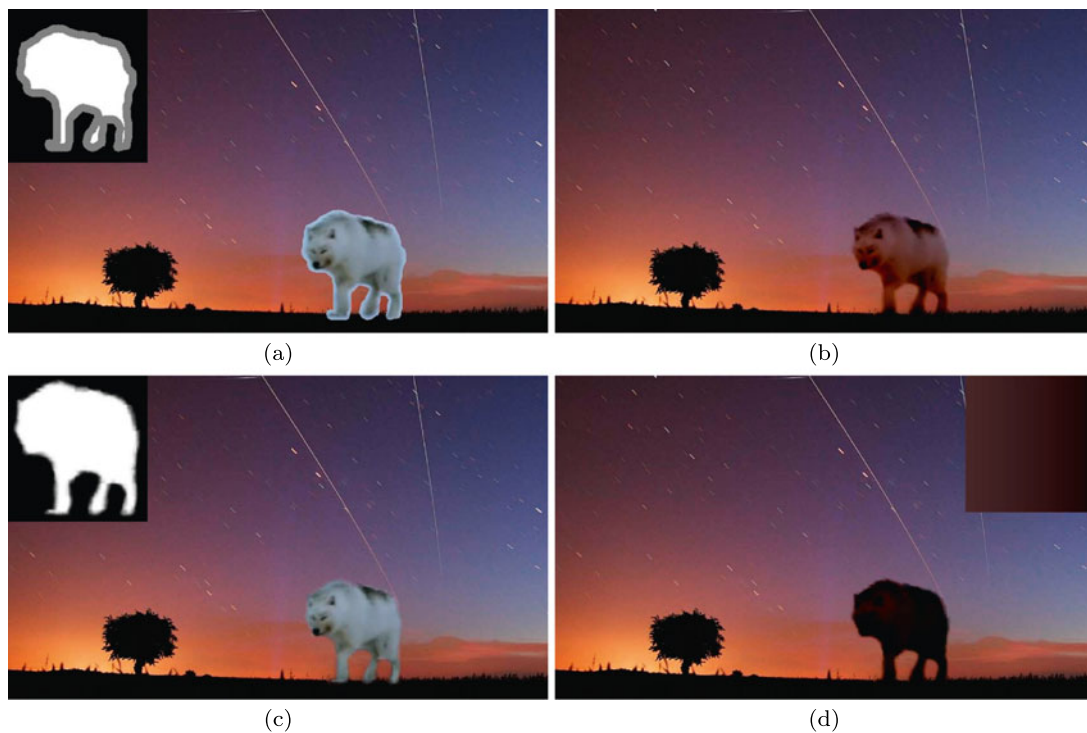
### 3.4 MVC cloning with hybrid boundary

Chen et al. [3] classified the pixels near the cloning boundary into two types:  $M_1$  consists of pixels where texture and color between the source and target are consistent, and  $M_2$  refers to the inconsistent boundary. In  $M_1$ , direct Poisson Blending is used by solving a Poisson equation with a Dirichlet boundary, while in  $M_2$ , the Poisson equation is solved with a Neumann boundary. Thus, the cloning result is

calculated by solving a Poisson equation with hybrid boundary, as shown in (9).

$$\Delta f' = \text{div } \mathbf{v} \text{ over } \Omega, \quad \text{with} \\ f'|_{M_1} = f^t \quad \text{and} \quad \nabla f'|_{M_2} = \nabla f_f^s. \quad (9)$$

In the previous sections, we assume that the boundary of source patch and target image are consistent, thus the MVC-based method can be applied directly in image cloning. However, when their boundaries are inconsistent, direct MVC-based approach will cause smudging and discoloration artifacts, see column (c) of Fig. 5. In [3], the Poisson blending with hybrid boundary involves solving a large linear system, which is memory and time consuming. Thus, we propose an interpolation method based on MVC with hybrid boundary to improve the efficiency. For the boundary in  $M_2$ , their first derivatives are known while their values are unknown. First, we set their values to be  $x_1, x_2, x_3, \dots$ . Then, we calculate the MVC interpolation of the points  $n_1, n_2, n_3, \dots$  which is on in the neighbor of  $M_2$ , see column (a) of Fig. 5 for details. Thus, the values of  $n_1, n_2, n_3, \dots$  are a linear combination of  $x_1, x_2, x_3, \dots$



**Fig. 4** Creating a new image by cloning the wolf into the darkness: (a) Source and target images (*trimap* is shown in *top-left corner*); (b) result of direct MVC Cloning [8]; (c) result of content-aware

cloning [7] (alpha matte is shown in *top-left corner*); (d) result using our method (reference image representing the global feature is shown in *top-right corner*)

Next, a linear system can be figured out according to the Neumann boundary in  $M_2$ , see (10).

$$\begin{aligned}
 x_i - \sum_{p \in M_1} w_i(p)(f^t(p) - g(p)) \\
 - \sum_{q \in M_2} w_i(q)(x_q - g(q)) = 0
 \end{aligned} \quad (10)$$

where  $w_i(p)$  and  $w_i(q)$  are the MVC interpolation weights for neighboring points  $n_1, n_2, n_3, \dots$  of  $x_i$  in  $M_1$  and  $M_2$ . This non-sparse linear system is small in size, which could be solved very fast by LU factorization. After calculating the values  $x_1, x_2, x_3, \dots$  on the boundary in  $M_2$ , direct MVC is applied to get the final result. See column (d) of Fig. 5 for MVC-based cloning results with hybrid boundary, the results show that our method can produce more natural results than previous gradient-based methods, such as MVC, content-aware image cloning. In addition, the time and memory costs are largely reduced compared with solving a large linear system in Poisson blending, which makes the cloning with hybrid boundary possible for practical use, e.g. image editing and video cloning.

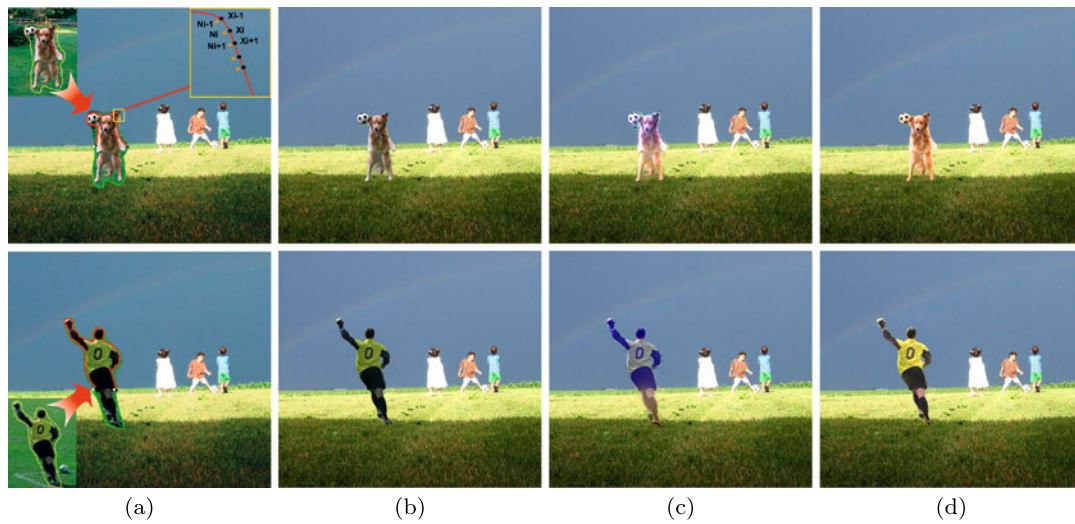
### 3.5 Application to video cloning

Given the fine performance of Environment-Sensitive image cloning, it is natural for us to apply it to video cloning. The

task of seamless video composition is usually divided into two steps: (1) Extracting the foreground by video cutout. (2) Blending it with the background naturally.

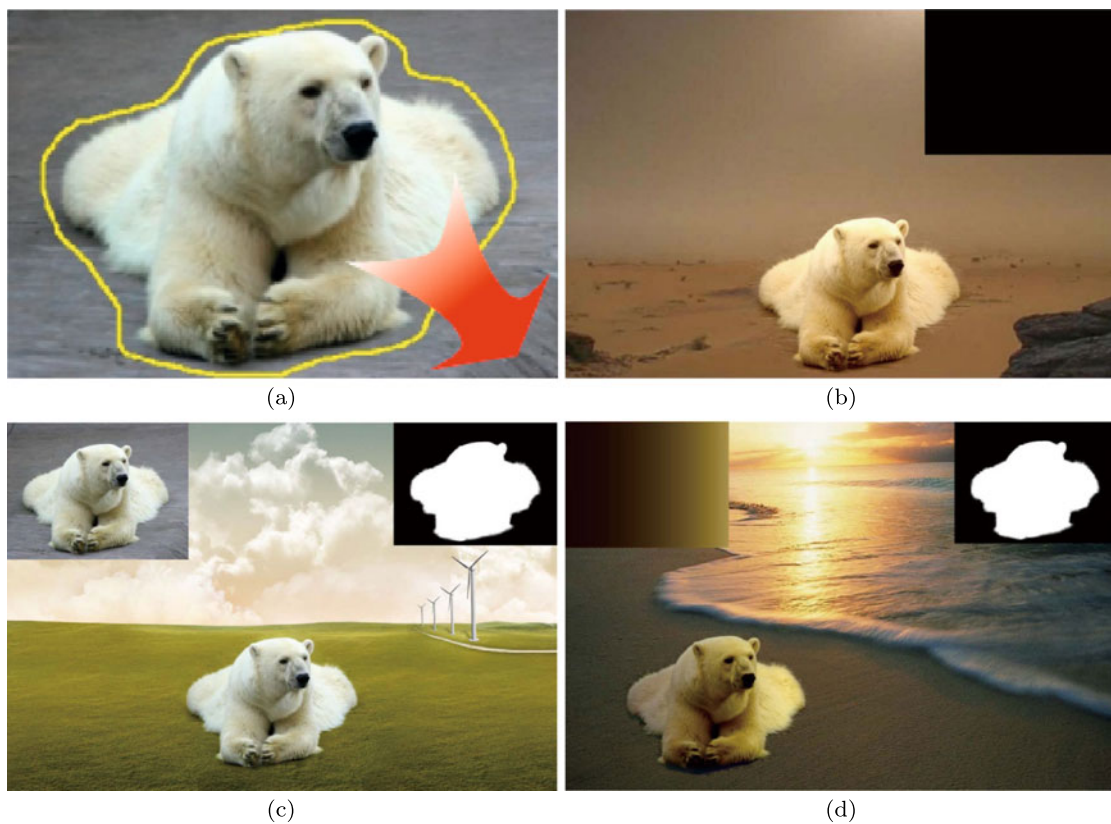
**Efficient video cutout** The extraction of foreground objects are based on the state-of-the-art Video SnapCut [2]. In order to reduce the time of user interaction, we modify the flow of Video SnapCut by replacing the matte with mask as output. Since matting is the most time-consuming and mainly depends on the foreground mask and source video, the matte of every frame can be calculated automatically after we get the mask sequence. We employ [14] to obtain high-quality matte. See Fig. 7 for the flow chart of our video cutout system in details.

**Video seamless cloning** After we have obtained the *trimap* and temporal coherent alpha matte of the source video, video cloning is degraded as the frame-by-frame image composition. For every frame, the outer boundary of the *trimap* is used as the cloning boundary, and the alpha matte can be used to control the diffusion of global and local features. By providing different reference image  $\xi$  and alpha matte according to the illumination feature of the target scene, our general image cloning model can be used to paste every frame into the target scene naturally. When considering the global illumination feature, we assume that it does not



**Fig. 5** MVC cloning result with hybrid boundary: The first and second row show a football player and a dog being pasted into the same grass land, respectively; column (a) shows the hybrid boundary (green

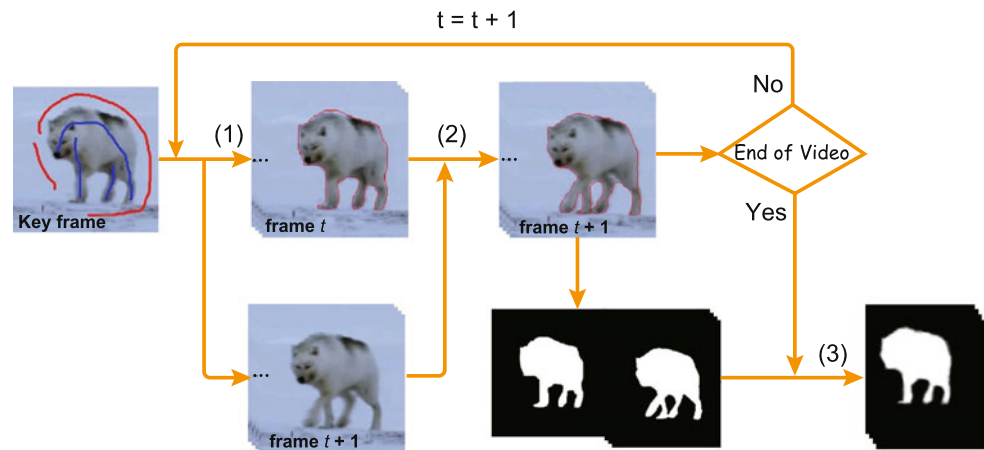
means the  $M_1$  boundary, and red means the  $M_2$  boundary); columns (b)–(d) are cloning results using content-aware [7], MVC [8] and hybrid MVC cloning, respectively



**Fig. 6** Visual appealing results produced by our cloning model: (a) Source image patch; (b) set  $\alpha = 0$  for all pixels in the cloning region, and the result is the same as the traditional MVC cloning [8]; (c) set  $\xi = g$ , and the result is the same as the content-aware cloning [7] (the alpha matte is shown in the top-left corner, while the top-right

corner is the source image (g)); (d) visual appealing cloning result which considers the global illumination feature (the alpha matte is shown in the top-left corner, while the top-right corner shows the reference image  $\xi$ )

**Fig. 7** The flow chart of our improved video cutout system



change during the video cloning process. Thus the global illumination feature of the target video can be extracted and represented by the method used in image cloning. As shown in Fig. 8, the appearance of reference images are similar in each frame, and their sizes are decided by the boundary box of the cloning area.

#### 4 Experimental results

In this section, we show the cloning results by our approach and make comparisons with previous methods. The experiments are performed in a desktop equipped with a 2.6 GHz of Intel Dual-Core CPU E5300, 4 GB memory and NVIDIA GeForce 9600 GT.

In Fig. 3, a bear is pasted into different target scenes featured with the sunset from the right and left side, respectively. The quality of cloning result would degrade when the global illumination feature is ignored. As shown in (d), (g) and (e), (h), the seamless but unnatural results are produced by MVC and content-aware cloning [7], respectively. Only (f) and (i) are satisfying results by our method which diffuses the global and local features into the cloning area.

Figure 4 shows a case that the illumination feature is indicated by the existing objects in the scene. The color appearance of a tree in the target scene indicates that the wolf to be cloned would appear to be as dark as the tree. Figure 4(a) is the source patch specified by the outer boundary of *trimap* and the target scene. Figures 4(b) and (c) are the results obtained using MVC-based and content-aware cloning, respectively. Although the two results are seamless, the former is unnatural when compared with the existing object (e.g. the tree), and the latter which preserves the patch's color appearance is unacceptable. The methods above do not work well because they ignore global illumination feature of the target scene. Figure 4(d) shows the satisfying result by our method.

When we extend the Environment-Sensitive image cloning to video cloning, a fine composition can also be achieved. In Fig. 8, there is a dominant illumination from the right side of the target scene, which is indicated by the vertical surface of the tree. As shown on column (a) of Fig. 8, cloning results of frame 7, 15, 21, 28 produced by MVC-based method are seamless but unnatural. When the global illumination is considered, the cloning results appear to be more acceptable, see column (b) of Fig. 8 for our result.

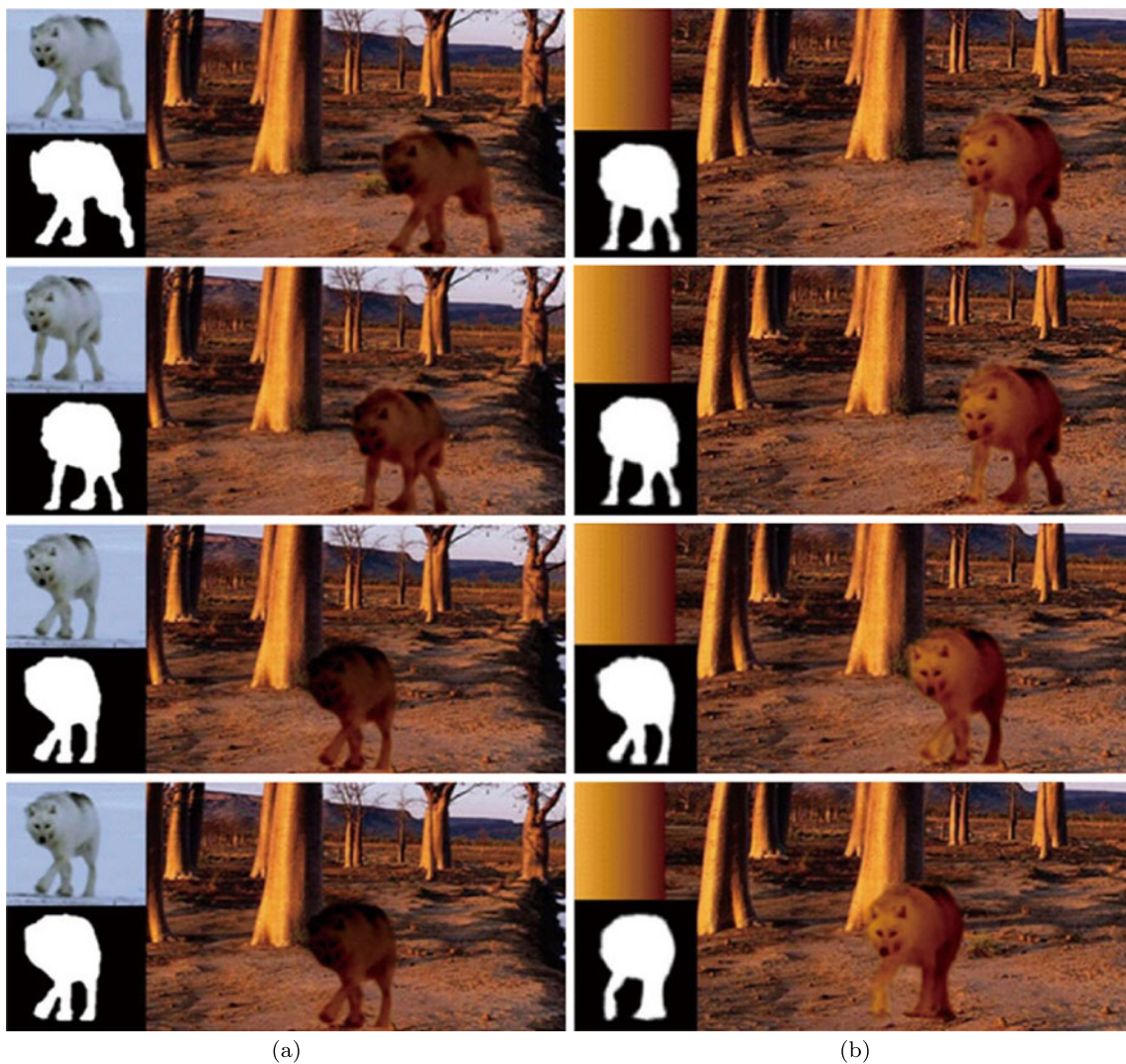
#### 5 Discussion

The implementation of our image cloning in (8) can be considered as a general model for image cloning as well. Specifically, when  $\alpha(x)$  is set to 1 or 0 in the cloning area, our method becomes original MVC-based cloning, whose target boundary is from the reference or target image, respectively. When  $\xi = g$ , the harmonic interpolant is expressed as

$$\sum_{i=0}^{n-1} \lambda_i(x) [(1 - \alpha(x))(f^*(P_i) - g(P_i))]. \quad (11)$$

The equation above is also an implementation of Ding's content-aware cloning [7]. Therefore, the general model can be used for image cloning by setting different  $\xi$  and  $\alpha$  to meet specific requirements. In Fig. 6, our image cloning model is used to paste a bear into different scenes by setting different values for  $\xi$  and  $\alpha$ . In (b), since the local and global features of the scene are compatible, thus the local feature has to be diffused into the bear, and we simply set  $\alpha(x) = 0$  to all pixels in the cloning area. In (c), we have to preserve the original appearance of the bear while prevent the diffusion of local feature, which is the same as content-aware [7]. Thus we simply set  $\xi = g$  and employ matting technique to meet this requirement. In (d), we create a reference image  $\xi$  according to global illumination feature of the target scene. Then, the global and local features are diffused





**Fig. 8** Seamless video cloning about a wolf: (a) Cloning results of frame 7, 15, 21, 28 using direct MVC cloning [8] (the source image is shown in *top-left corner* and foreground mask is in *bottom-left corner*);

(b) cloning results of frame 7, 15, 21, 28 using our method (new target image is in *top-left corner* and alpha matte is in *bottom-left corner*)

into the cloned patch under the control of  $\alpha$ . As can be seen, the final result is visual appealing.

## 6 Conclusions

We have given an Environment-Sensitive image cloning technique which takes into account the global and local features of the target scene. A reference image is created to represent the global visual feature of the target image, which can be further diffused into the cloning area. Matting technique is used to control the diffusion of global and local features of the target image into the cloning area. This can remove the visible seam while ensuring that the color appearance of the cloning region is in accord with the target image. In addition, an efficient solution using the MVC frame-

work is proposed to composite images with hybrid boundary. Moreover, based on our environment-sensitive cloning, we construct a general model for image composition to adjust the cloning result by setting different parameters. We further apply our image cloning method to video cloning. Experimental results show the advantages of our method over the previous approaches.

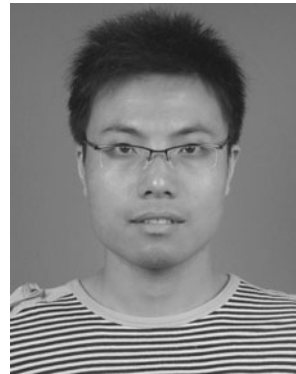
In the future, we would like to explore more effective methods to extract global visual features and create more complex reference images, so as to produce more visual pleasing results. In addition, roughly analyzing the geometric character of the foreground objects is another aspect to improve the cloning result.

**Acknowledgements** We thank all anonymous reviewers for their valuable comments. This work is supported by the National Basic Re-

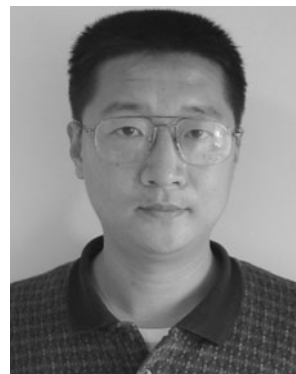
search Program (No. 2011CB302205) of China and National High-Tech Research and Development Program (No. 2009AA01Z330) of China.

## References

1. Armstrong, C.J., Price, B.L., Barrett, W.A.: Interactive segmentation of image volumes with live surface. *Comput. Graph.* **31**(2), 212–229 (2007)
2. Bai, X., Wang, J., Simons, D., Sapiro, G.: Video snapcut: robust video object cutout using localized classifiers. *ACM Trans. Graph.* **28**(3), 70 (2009)
3. Chen, T., Cheng, M.M., Tan, P., Shamir, A., Hu, S.M.: Sketch2photo: Internet image montage. *ACM Trans. Graph.* **28**(5), 124 (2009)
4. Cheng, M.M., Zhang, F.L., Mitra, N.J., Huang, X., Hu, S.M.: Repfinder: finding approximately repeated scene elements for image editing. *ACM Trans. Graph.* **29**(4), 83 (2010)
5. Chuang, Y.Y., Curless, B., Salesin, D.H., Szeliski, R.: A Bayesian approach to digital matting. In: *Proceedings of IEEE CVPR 2001*, vol. 2, pp. 264–271. IEEE Computer Society, Los Alamitos (2001)
6. Chuang, Y.Y., Agarwala, A., Curless, B., Salesin, D.H., Szeliski, R.: Video matting of complex scenes. *ACM Trans. Graph.* **21**(3), 243–248 (2002). Special Issue of the SIGGRAPH 2002 Proceedings
7. Ding, M., Tong, R.F.: Content-aware copying and pasting in images. *Vis. Comput.* **26**(6–8), 721–729 (2010)
8. Farbman, Z., Hoffer, G., Lipman, Y., Cohen-Or, D., Lischinski, D.: Coordinates for instant image cloning. *ACM Trans. Graph.* **28**(3), 67 (2009)
9. Gastal, E.S.L., Oliveira, M.M.: Shared sampling for real-time alpha matting. *Comput. Graph. Forum* **29**(2), 575–584 (2010). *Proceedings of Eurographics*
10. Jia, J., Sun, J., Tang, C.K., Shum, H.Y.: Drag-and-drop pasting. *ACM Trans. Graph.* **25**(3), 631–637 (2006)
11. Lalonde, J.F., Efros, A.A., Narasimhan, S.G.: Estimating natural illumination from a single outdoor image. In: *IEEE International Conference on Computer Vision* (2009)
12. Lalonde, J.F., Hoiem, D., Efros, A.A., Rother, C., Winn, J.M., Criminisi, A.: Photo clip art. *ACM Trans. Graph.* **26**(3), 3 (2007)
13. Land, E.H., McCann, J.: Lightness and retinex theory. *J. Opt. Soc. Am.* **61**, 1–11 (1971)
14. Levin, A., Lischinski, D., Weiss, Y.: A closed-form solution to natural image matting. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(2), 228–242 (2008)
15. Li, Y., Ju, T., Hu, S.M.: Instant propagation of sparse edits on images and videos. *Comput. Graph. Forum* **29**(7), 2049–2054 (2010)
16. Pérez, P., Gangnet, M., Blake, A.: Poisson image editing. *ACM Trans. Graph.* **22**(3), 313–318 (2003)
17. Ruzon, M.A., Tomasi, C.: Alpha estimation in natural images. In: *CVPR*, pp. 1018–1025 (2000)
18. Wang, J., Cohen, M.F.: An iterative optimization approach for unified image segmentation and matting. In: *ICCV*, pp. 936–943 (2005)
19. Wang, H., Raskar, R., Ahuja, N.: Seamless video editing. In: *ICPR* (3), pp. 858–861 (2004)
20. Wang, J., Agrawala, M., Cohen, M.F.: Soft scissors: an interactive tool for realtime high quality matting. *ACM Trans. Graph.* **26**(3), 9 (2007)
21. Wang, J., Cohen, M.F.: Image and video matting: a survey. *Found. Trends Comput. Graph. Vis.* **3**(2), 97–175 (2007)
22. Xie, Z.F., Shen, Y., Ma, L., Chen, Z.: Seamless video composition using optimized mean-value cloning. *Vis. Comput.* **26**(6–8), 1123–1134 (2010)
23. Xu, K., Li, Y., Ju, T., Hu, S.M., Liu, T.Q.: Efficient affinity-based edit propagation using k-d tree. *ACM Trans. Graph.* **28**(5), 118 (2009)



**Yun Zhang** is a Ph.D. candidate in Department of Computer Science, Zhejiang University, China. He received his B.E. and M.A. from Hangzhou Dianzi University, China in 2006 and 2009. His research interests include image and video processing and computer graphics.



**Ruofeng Tong** is a professor in Department of Computer Science, Zhejiang University, China. He received his B.E. from Fudan University, China in 1991, and a Ph.D. from Zhejiang University, China in 1996. He was a visiting researcher in Faculty of Engineering, Hiroshima University, Japan in 2000. His research interests include image and video processing, computer graphics, and computer animation.