Automatic Foreground Extraction of Head Shoulder Images

Paper No.: 829

Abstract

Most existing techniques of extracting foreground work only in interactive mode. This paper introduces a novel algorithm of automatic foreground extraction for special object, and verifies its effectiveness with head shoulder pictures. The main contribution of our idea is to make the most of the prior knowledge to constrain strongly the processing of foreground extraction. For human head shoulder images, we first detect face and a few facial features, which help estimate an approximate mask covering the interesting region, based on the Candide face model (which is the subset of MPEG4 face standard model); the algorithm then extracts the hard edge of foreground from the specified area, using an iterative graph cut method incorporated with an improved Gaussian Mixture Model. To generate accurate soft edges, a Bayes matting is employed. The whole process is fully automatic. Experimental results demonstrate that our algorithm is both robust and efficient.

keywords: *Foreground extraction; Graph cut; matting*

1. Introduction

Image segmentation has been a hot topic and difficult problem in computer vision for many years. Early research focused on fast and precise interactive segmentation tools, thereinto, Intelligent Scissors [6], Snakes [4] and Level sets [5] are several typical methods, which present hard edges. Then, matting techniques were developed to get soft edges with one transparent alpha mask, Poission [9] and Bayes [8] matting are the two representative algorithms. Nevertheless, all above works request user interaction; recently, some researchers attempted to enhance the automatization [1, 2, 3]. In [3], an intuitive user interface is suggested, but the interaction is still complex; [1, 2] present the most automatic technique on image cutout update. [2] need only draw a rectangle surrounding the desired object. But, the algorithm is not robust enough since it can not present some precise region to reduce the unwanted samples on the background and foreground.

Indeed, automatic foreground extraction without user intervention is a big challenge for image matting. It seems impossible to develop a universal algorithm to recognize the approximate outline of the interesting object from complex background, because the shape of objects in nature vary greatly, such as cars, human, flowers and so on. To make the algorithm more automatic, we should take some knowledge of the foreground objects into account. Actually, for many special objects, there are some common characteristics which are valuable and helpful to locate objects' outline. In fact many detection techniques for special objects have been proposed, such as car, face, license plate of car, text, and so on. Those detection results will facilitate automatic image segmentation. Based on above analysis, we present a novel framework of fully automatic foreground extraction for special objects. Our paper focuses on making the most of the prior knowledge to select more precise sample set of the background and foreground, remove unwanted pieces and retrieve lost patches.

Our algorithm is divided into three stages. First, we detect the objects' probable outline or locate some obvious features, which used to help create a rough mask covering the interesting region. Secondly, with the prior knowledge, an iterative Graph cut algorithm based on improved Gauss Mixture Model (GMM) is employed to estimate the foreground over the mask and generate the hard edges. Thirdly, along a narrow strip around the hard edges, Bayes matting is used to refine the cutout boundary and generate the last soft edges.

Human face and body are the most familiar things in our life, whose extraction is significant and meaningful. So we select the head shoulder pictures as examples to validate our algorithm. Its prior knowledge can be represented as facial organ, which can be obtained with face detection and features locating.

The main contribution of this paper is that we introduce a novel, robust algorithm to extract foreground automatically for special objects, in which we make full use of the prior knowledge as strong constraint. Matting is also used to refine edges and generate alpha mask. This work is of great significance for those embedding device without interactive tools, such as digital camera and mobile.

The remainder is organized as follows. In section 2, a brief review and analysis about the related work is presented. Section 3 describes the algorithm of automatic foreground extraction for head shoulder photos in detail. Section 4 demonstrates some experimental results and Section 5 concludes the whole paper and highlights future work.

2. Related Work

For automatic image cutout, there are few achievements up to now. We will describe briefly and compare several representative interactive image cutout techniques, which focus on present simple and convenient interaction. In addition, GMM will be addressed briefly which is employed in our improved Graph cut algrithm.

Image cutout Several early image segmentation methods have been applied in practical products, such as Snakes [4], Level sets [5] and Intelligent Scissors [6], all of which are interactive tools. During their application, user first define an initial contour; then driven by different forms of energy minimization, these contours can be refined and snapped to the image edges iteratively. The main disadvantage of above methods is that the final segmentation result may depend on the initialization heavily, especially in camouflage; excessive interaction also limits their application.

Recently, a powerful image cutout technique, Graph cut [1], is presented to enhance the automation of segmentation, which is addressed in monochrome image firstly. Graph cut poses image cutout as a binary labelling problem, which means that each pixel p belongs to either foreground or background. Define L_p as the transparency of p, if p belongs to foreground, denote it as $Lp = 1$; otherwise $Lp = 0$. Guided by the observed foreground and background greylevel histograms, the solution L, i.e. a good segmentation, can be obtained by minimizing a Gibbs energy $E(L)$ with the form:

$$E(L) = U(p, L_p) + V(p, L_p) \tag{1}$$

where the first term represents penalty energy for pixel labelling and the latter term is the smoothing term indicating the interaction between neighboring pixels. Usually $U(p, L_p)$ and $V(p, L_p)$ can be defined as follows:

$$U(p, L_p) = \lambda \cdot \sum_{p \in P} -lnh(p; L_p), \tag{2}$$

$$V(p, L_p) = \sum_{(p,q) \in N} \frac{exp(-(I_p - I_q)^2/2\sigma^2) \cdot |L_p - L_q|}{dist(p, q)} \tag{3}$$

in which λ is in fact a weight between the penalty energy and the smoothing term, $(p, q) \in N$ means p, q are neighboring pixels, h is the color distribution, I_p is the grey value of pixel p, σ is a constant and $dist(p, q)$ is the Euclidean distance between p and q. Minimization of $E(L)$ is done by using a standard minimum cut algorithm. More details about the Graph cut algorithm can be found in [1].

GrabCut [2] extended Boykov's algorithm on several aspects to improve the efficiency and reduce user interaction. Firstly, it replaces constructing a large of color histograms with GMM to achieve more accurate results, which is more reasonable in principle also. Secondly, GrabCut adopts an iterative Graph cut procedure to substitute for the one-shot minimum cut estimation, which can escalate precision. It needs only one rectangle dragged surrounding the desired object for foreground extraction, which is the most automatic algorithm for extracting foreground under complex background up to now. However, this algorithm is not robust, since it can not select precise samples set for foreground/background to construct exact GMM in many cases; then only rough, even partly false GMM is generated.

Lazy Snapping [3] is another fine Graph cut based cutout algorithm. It separates coarse and fine scale processing, making object specification and detailed adjustment easy. In the coarse process, the image is clustered adaptively and Graph cut is applied to these clusters to generate an initial segmentation; while in the fine process, a set of user interface tools is designed to provide flexible control and editing. To obtain satisfactory results, the exigent interaction is the unavoidable disadvantage.

Above approaches mainly focus on the hard segmentation, which can not get much precise and smooth edge information. Their results can not be fused into another image smoothly. A few works have been addressed in the soft segmentation [7, 8, 9], which endows the pixel around the boundary with continuous alpha values. Usually, for these approaches, a user-supplied trimap $T = \{T_F; T_U; T_B\}$, is needed, where T_F is the user specified foreground, T_B is background and T_U is the the unknown region whose alpha values are to be solved. Bayes matting [8] models the color distributions with oriented Gaussian covariance, and relies on a Bayesian approach to solve the matting problem. Poisson matting [9] formulates the problem as solving Poisson equations with the matte gradient field.

Gaussian Mixture Model (GMM): is a type of probability density composing of a set of Gaussian models. The GMM with K Gaussian models is:

$$p(\omega|x) = \sum_{k=1}^{K} \alpha_k G(x; \mu_k, \sigma_k) \tag{4}$$

where $\alpha_1, ..., \alpha_k$ is the mixing proportions satisfying $\Sigma_{k=1}^{K} \alpha_k = 1$, μ_k, σ_k is the mean value and covariance of kth Gaussian model.

A commonly used approach for determining the parameters of GMM is the Expectation-Maximization(EM) algorithm [10].

3. Head Shoulder Cutout Algorithm

The process of automatically extracting foreground of head shoulder pictures is divided into three steps: make a rough mask with the face detection and feature location techniques; extract the hard edges of the foreground with the constraints of the mask by Graph cut; refine the result and generate soft edges with matting.

3.1 Face and feature detection

Face and facial feature detection have been studied for many years, and many detection techniques have been presented, [15, 16, 18]. A detailed comparison and survey on face detection algorithms can be found in [15]. Among the face detection methods, those algorithms based on learning have demonstrated excellent results. Recently, P. Viola [18] presented an efficient and robust method based on AdaBoost, with an image representation called integral image allowing fast feature evaluation. We adopt an improved Adaboost algorithm to search the face candidate in head shoulder image.

In order to enhance the detection ratio and performance of Adaboost algorithm, we make two improvements. First, because its speed is correlated directly to search space, an adaptive Gauss mixture skin color model is employed to ignore those non-skin regions, and canny filter is used to reduce those chaos regions. Second, Adaboost algorithm exists a drawback, that is, when the number of features exceeds 200, the distribution of face and non face classes in Harr-like features space almost completely overlap in later stages of the cascade training. In order to solve this problem and improve the detection performance, PCA feature is introduced in the later stages of cascaded training. Adaboost with PCA can get much higher detection ratio than that without PCA at the same false alarm ratio.

Further, based on the detected face, we also employ Adaboost algorithm to locate eye corners and mouth corners, and apply VPF [13] algorithm to locate jaw and jowl feature points. The whole process abides to the idea of LFA [14], which can enhance the efficiency and increase detection ratio. In general, we detect 9 feature points: eye corners, mouth corners, left/right jowl and jaw.

3.2 Face mask

With the help of the the detected face and features, we can create a mask to define the interesting region of the image, which indicates the approximate positions of the head and shoulder. This mask will strongly constrain the whole Graph cut process to emphasize foreground, remove fractional pieces and unwanted patches, and retrieve the lost foreground.

Candide is a wire-frame face model to define a basic facial structure, which has been popular in video coding research for many years [19]. The new variant of this model is also compliant to MPEG-4 Face Animation. We adopt

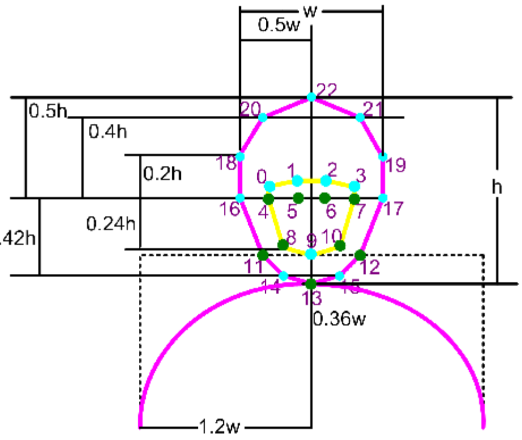


Figure 1: The head and shoulder mask. The green points are automatically detected and the blue ones are estimated with Candide model; Some length proportions are also labelled, wherein h, w represent the length and width of the face respectively.

the frontal projection of Candide to assist in constructing the face mask, as shown in Fig.1. The green points denote the detected features and the blue ones are estimated with Candide. The region surrounded by the pink polygon demonstrates the approximate face area and the half ellipse adhering to the face mask denotes the shoulder position. Besides, the proportions of all parts are also fixed.

In Fig. 1, the area covered by face mask and the half ellipse is deemed as the estimated initial foreground served for the further Graph cut, while the rest region is the background; the initial samples selection of background/foreground is more accurate than previous semi-automatic approaches.

3.3 Color distribution

For describing color distribution, GMM is a popular and typical model which has been widely applied. We found on it exploring our model, named multi-GMM, to arrive more rapid and robust effect. If we employ a single GMM directly, precise cut results can not be obtained in many cases. For a GMM, the final cut effects and efficiency are sensitive to three main facets: the number of color clusters, the initialization method of GMM and fitting degree between generated GMM and actual color datum. Firstly, we employ the algorithm in [17] to determine proper number of color cluster. If specify forcibly the number as most of the previous works, poor cut results may be produced in a few cases. Secondly, there are three class methods: splitter [20], k-means [21] and LBG [22] to initialize GMM. We select LBG algorithm, which can make more precise initial

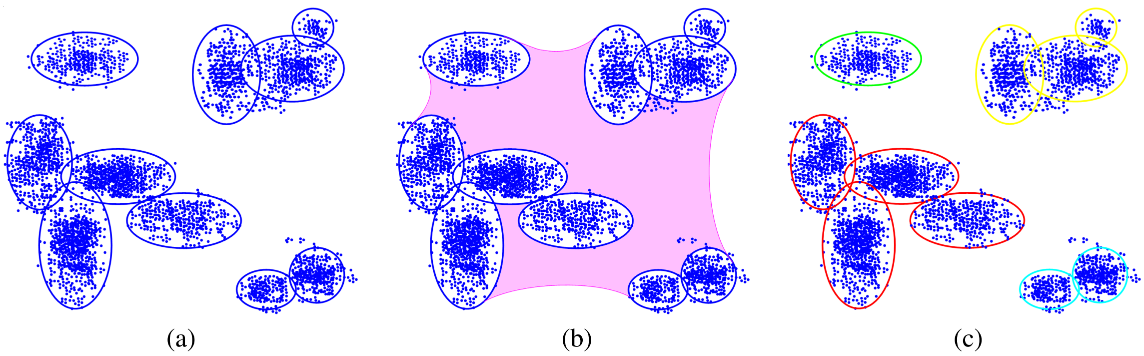


Figure 2: Multi-GMM (2-D case). (a) Color clustering; (b) The color is modelled with a single GMM; (c) The color is modelled with multi-GMM, each of which is represented with a different color.

value with rationally more cost. It is iteratively split and re-estimated to yield an initial GMM. Thirdly, without doubt, the more closely the color distribution model fits the color datum, the more precise cut results will be. This is a key factor for the most difficult case of background and foreground being much similar. We apply multi-GMM to generate more tight color distribution description. As shown in Fig. 2(b), if we model all data with a single GMM, it may be inaccurate (pink region are false parts). So, we divide it into multi GMMs. According to the correlation degree among every color clusters, we divide the whole space into several subspaces, and each subspace is regarded as one absolute GMM. Fig. 2 (c) shows the multi-GMM, four GMMs are labelled with four color ellipses separately. Still we apply traditional EM algorithm [11] to model our GMMs. Either the estimated foreground or the estimated background corresponds to a multi-GMM. Actually, smaller number of clusters will be produced based on our dividing sub-region idea narrated in the next paragraph, which can also increase efficiency.

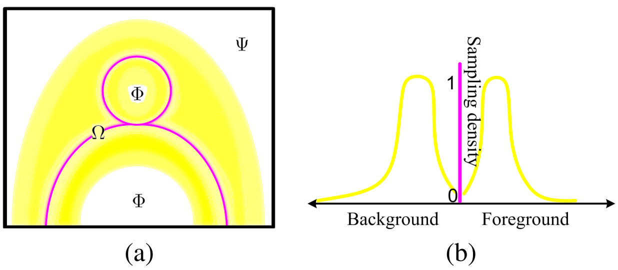


Figure 3: Pixel sampling principles. (a) The thick yellow and light yellow area represent the high and low sampling area separately, while the blank means non-sampling; (b) visualizes the sample function.

For sampling, selecting all pixels within mask to construct the multi-GMM will result in great mistake since the face mask and half ellipse are just a rough estimation of the head shoulder positions. To reduce the affection of above problem and decrease the computational complexity for constructing multi-GMM, we adopt the sampling principles as follows. First, in general cases, the central area of the mask and half ellipse should belong to foreground doubtless, which is deemed as foreground without sampling (Φ in Fig. 3 (a)). This area is almost worthless for the segmentation, on the contrary, it may bring some false cut patches sometimes in principle; moreover, sampling this region will cost highly for constructing the multi-GMM. Secondly, the region near the boundary of the mask and half ellipse is the unknown area (Ω in Fig. 3 (a)); the narrow strip near the boundary will be sampled sparsely while gradually becoming dense away from the boundary; the sampling ratio is lowest on boundary. Thirdly, similar to the central region, the region (Ψ in Fig. 3 (a)) near the image border is the background certainly, so its sampling ratio should be low or zero.

We adopt the probability density function of F-distribution to guide us sampling the foreground/background:

$$P_F = \frac{\Gamma((m+n)/2)}{\Gamma(n/2)\Gamma(m/2)} n^{\frac{n}{2}} m^{\frac{m}{2}} \frac{x^{n/2-1}}{(nx+m)^{(m+n)/2}} \quad (5)$$

where Γ is the Gamma distribution, the parameters m, n determine the sample principle (we set m, n to 10 in our experiment). The gradual changing of sampling ratio is visualized as Fig 3. In (a) the sample density of the thick yellow area represents high sampling ratio and the light yellow means lower, whereas the blank means non-sampling; (b) visualizes the sampling function, reflecting the relation between the sample ratio and the distance of a certain area to the pink boundary.

In addition, we divide the whole interest object into multiple sub-regions, so that we can reduce color cluster number, generate more fitting GMM, and improve efficiency as well. For head shoulder photos, we divide them into two sub-regions of head and shoulder, in which shoulder mask can be adaptively created based on the head clustered result.

For other special objects, we can divide different number of sub-regions according to its physical structure.

3.4 Image cutout

With the multi-GMM, an improved Graph cut algorithm is employed to cut out the head and shoulder from the pictures.

For the description convenience, as in Fig. 3 (a), we define the pink curve C as the contour consisting of the outer edge of the face mask and its adhering half ellipse. We represent the sampled pixel set out of C as P_b, the interior set as P_f. Obviously, for the pixel belonging to P_b , the farther from C, the more probable it is background; for the sampled pixel of P_f , the case is similar. This fact is imposed in Graph cut by appending prior weight function to the penalty energy term of the Graph cut objective function.

For a pixel p, $dis(p)$ denotes the Euclidean distance from p to C. If p belongs to P_f, define $D(p) = dis(p)$ as the distance from p to the contour; otherwise, define a negative, $D(p) = -dis(p)$ as its oriented distance to C. We adopt a normalized function $W(p, L_p)$ as the prior coefficient with the form:

$$W(p, L_p) = \begin{cases} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{p} e^{-\frac{x^2}{2}} dx + \frac{1}{2} & : \quad L_p = 0 \\ 1 & : \quad L_p = 1. \end{cases}$$

The Gibbs energy of (1) is modified as:

$$E(L) = W(p, L_p) \cdot U(p, L_p, G_p) + V(p, L_p) \qquad (6)$$

where $W(p, L_p)$ acts as the intensity to emphasize the constraint of prior knowledge and is valued as the displacement of the distribution function of the standard normal distribution for $L_p = 0$. It takes effect on the background energy or foreground energy. For background energy, if p belongs to P_f, its background energy is reduced, otherwise increased.

$U(p, L_p, G_p)$ derived from $U(p, L_p)$ of (1) is the new penalty energy incorporated with GMM by substituting the pixel's weighted Gaussian distributions: G_p for the color distribution in monochrome image. It can be formulated as:

$$U(p, L_p, G_p) = \lambda \cdot \sum_{p \in P} [-ln(\bigodot_{i=1}^{n} \alpha_i(p) \sum_{j=1}^{l_i} \beta_{ij} G(p, \mu_{ij}, \sigma_{ij}))] \qquad (7)$$

in which λ inherited from (1). n represents the number of GMM and l_i means the number of Gaussians contained in ith GMM. The symbol $\bigodot$ represents whether p belongs to ith GMM. In our color multi-GMM distribution, each pixel p only belongs to a GMM in general, assumed as kth ($k \in [1, n]$), then $\alpha_k(p)$ is set to 1, others set to zero. β_j is the coefficient of jth Gaussian's probability weight among ith GMM and is computed by dividing the sample number in jth Gaussian by the pixel number contained in ith GMM.

We adopt an iterative method to solve $E(L)$. It starts with minimizing $E(L)$ with the standard minimal cut algorithm, and regards the cutout result as the initial estimation of the foreground/background for the next iteration. After every iteration, the multi-GMM is applied again to model the color data of the new estimated foreground/background. Since after every iteration the estimated foreground/background becomes more accurate, the iteration is convergent.

In terms of our new form of Gibbs energy, for every iteration, the pixels on estimated foreground/background is re-sampled and clusted by LBG algorithm, while EM optimization is performed iteratively. The rules for computing $W(p, L_p)$ holds unchanged, so the distance $D(p)$ for every sample is computed only once.

3.5 Matting

Image matting used to generate soft cutout edges, especially useful for segmenting transparent objects, such as the hair in this paper. We perform Bayes matting along a trimap around the hard edges generated with above Graph cut. The unknown area of trimap is generated by dilating along the hard edges with a constant width (we set 12 pixels in our experiments), while interior and exterior regions about the unknown region are labelled as foreground and background separately.

Table 1 summarizes the whole process of our image cutout algorithm in detail. Note that, to enhance the efficiency of our algorithm, Graph cut can be implemented on the down-sampled image, whereas the matting algorithm is performed on the original image, which doesn't basically affect the last results in principle.

4. Experimental Results

We have tested our algorithm with 263 normal upright head shoulder pictures with over 70% success ratio on an Intel Pentium IV 2.4GHz PC with 512MB main memory under the Windows XP operating system. The detection costs 1-2 seconds, cutout stage costs 1-3 seconds and matting takes 1-3 seconds.

Fig. 4(b), (e) show the poor cutout results without the constraint of prior weight function $W(p, L_p)$ while solving the objective function $E(L)$. The reason lies in that the color distribution of foreground is similar to its nearby background. (c), (f) are the right cutout results by taking the constraint of prior knowledge into account.

Fig. 5 demonstrates the iterative process of solving $E(L)$. Here we give the cutout result of the head part step by step as (b), (c) and (d). On the other hand, although the facial detection result (a) is not fully exact inducing inaccurate mask, our algorithm can generate proper cutout (e) as well.

Table 1. Algorithm: Automatic image cutout

Step 1 Detect face and a few facial features.
Step 2 Estimate a face mask covering the interesting region.
Step 3 Sample pixels on the estimated foreground/background.
Step 4 Model the color distributions of the estimated foreground/background with multi-GMM:
  Step 4.1 Cluster color with the algorithm in [16] to determine the total number of color clusters;
  Step 4.2 Determine the number of GMMs according to the distance and overlapping between these clusters;
  Step 4.3 Construct every GMM:
    Step 4.3.1 Initial GMM with LBG algorithm;
    Step 4.3.2 Optimize GMM with EM algorithm.
Step 5 Cut out head shoulder with Graph cut:
  Step 5.1 Compute oriented distance $D(p)$ for each sampled pixel and compute $W(p, L_p)$;
  Step 5.2 Compute E(L) and solve it with the minimum cut algorithm to get a new estimated foreground/background;
  Step 5.3 Repeat to step 3, until convergence.
Step 6 Apply Bayes matting to generate soft edges.

Fig. 6 shows the matting effect of the hair. Fig. 7 presents some other examples. Fig. 8 composes a cutout result with a different background smoothly.

A few of the above pictures are fetched from the papers: [2] and [23].

5. Conclusions and Future Work

We have presented a novel approach to automatically extract special object and verified its effectiveness with head shoulder pictures. The key point lies in how to make full use of some prior knowledge of the special object to estimate the interesting region, model accurately the color distribution of the image foreground/background and implement a more precise, robust and rapid Graph cut algorithm with prior knowledge. The main advantage of this approach is that it breaks away from the user control without any manual interaction.

Although initial experiments generated some encouraging results, our approach is not yet robust enough to handle all cases. One main problem of the current technique is that our current face detection algorithm has not considered the cases of rotation face and side lighting photos, which will affect the cutout result sometimes. How to resolve this problem and abide this tolerance is a future work, as it is the common problem in the research area of face detection. Besides, the efficiency is not high yet, the extraction of any size image will cost from 2 to 5 seconds in general. Our future works include the following aspects. Part of the processing, for instance, the GMM construction or Graph cut, can be ported to the GPU (Graphics Processing Unit) that may realize the real-time cutout. We are also doing some research on extending our algorithm to extract the full body. In addition, similar to the head shoulder pictures, the automatic extraction of other special object, such as cars, trees, animals, etc., will be realized.

References

[1] Y. Boykov, and M.-P. Jolly, "Interactive Graph cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images. " In *IEEE International Conference on Computer Vision*, pp. 105-112. 2001.

[2] C. Rother, V. Kolmogorov and A. Blake, "Grabcut - Interactive Foreground Extraction Using Iterated Graph cuts." In *ACM SIGGRAPH 2004*, pp. 309-314. 2004.

[3] Y. Li, J. Sun, C. Tang and H. Shum, "Lazy Snapping." In *ACM SIGGRAPH 2004*, pp. 303-308. 2004.

[4] M. Kass, A. Witkin, and D. Terzolpoulos, "Snakes: Active Contour Models," *International Journal of Computer Vision*, Vol. 2, pp. 321-331, 1988.

[5] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic Active Contours." In *IEEE International Conference on Computer Vision*. pp. 694-699. 1995.

[6] E. N. Mortensen, and W. A. Barrett, "Intelligent Scissors for Image Composition." In *ACM SIGGRAPH95*, pp. 191-198. 1995.

[7] COREL Corporation. *Knockout user guide.* 2002.

[8] Y.-Y. Chuang, B. Curless, D. Salesin, and R. Szeliski, "A Bayesian Approach to Digital Matting." In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 264-271, 2001.

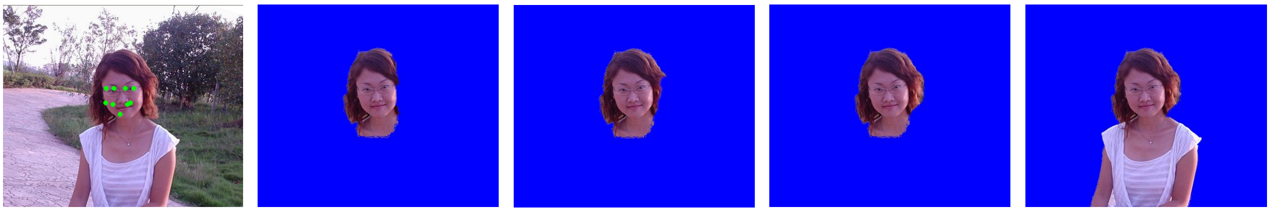
[9] J. Sun, J. Jia, C. Tang and H. Shum, "Poisson Matting." In *ACM SIGGRAPH 2004*, pp. 315-321. 2004.

[10] N. Vasconcelos and A. Lippman, "Embedded Mixture Modeling for Efficient Probabilistic Content-Based Indexing and Retrieval". In *Proc. of SPIE Conf. on Multimedia Storage and Archiving Systems III*, Boston, 1998.

[11] G. McLachlan and T. Krishnan, "The EM Algorithm and Extensions", Wiley Series in Probability and Statisticcs, John Weley & Sons.

(a) (b) (c) (d) (e) (f)

Figure 4: The constraint of prior knowledge. (a), (d) are the original images; (b), (e) are the cutout results without the constraint of $W(p, L_p)$, (b) loses a few pieces and (e) generates some unwanted patches; (c) (f) are the results taking account of $W(p, L_p)$.



(a) (b) (c) (d) (e)

Figure 5: The procedure of iteration. (a) The original image with detected feature points (the green points); (b), (c) and (d) show the iterative results of head; (e) The last cutout result.

[12] Z. B. Lai, P. Gao and T. wang et al., "Comparison on Bayesian YING-YANG Theory Based Clustering Number Selection Criterion with Information Theoretical Criteria", In *IEEE International Joint Conference on Neural Networks*, Anchorage, USA, Vol. 1, pp. 725-729, 1985.

[13] X. Geng, X. P. Zhong, X. M. Zhou, S. P. Sun and Z. H. Zhou, "Refining Eye Location Using VPF for Face Detection", In *Proc. of the 3rd Conference of Sinobiometrics*, Xi'an China, pp. 25-28, 2002.

[14] E. D. Mandel and P. S. Penev, "Facial Feature Tracking and Pose Estimation in Video Sequences by Factorial Coding of the Low-Dimensional Entropy Manifolds due to the Partial Symmetrie s of Faces", In *Proc. 25th IEEE Int'l Conf. Acoustics, Speech and Signal Processing (ICASSP 2000)*, Vol. 4, pp. 2345-2348, 2000.

[15] M. H. Yang, D. J. Kriegman and N. Ahuja, "Detecting Faces in Images: A Survey". In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24(1), pp. 34-58, 2002.

[16] T. S. Chua, Y. L. Zhao, M. S. Kankanhalli, "Detection of human faces in a compressed domain for video stratification", The Visual Computer. Vol. 18: 121-133, 2002.

[17] P. Gao and M. R. Lyu, "A Study on Color Space Selection for Determining Image Segmentation Region Number", In *Proceedings of the 2000 International Conference on Artificail Intelligence*, Monte Carlo Resort, Las Vegas, Nevada, USA, June 26-29, Vol 3, pp. 1127-1132, 2000.

[18] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", In Proc. *IEEE Conf. on Computer Vision and Pattern Recognition*, Kauai, Hawaii, USA, 2001.

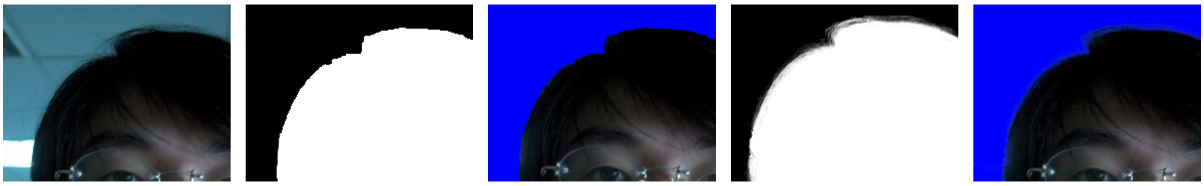
[19] J. Ahlberg, "Candide-3 – an Updated Parameterized Face," *Technical Report LiTH-ISY-R-2326*, Linkping University, Sweden, 2001.

[20] A. Senior, R-L Hsu, M. A. Mottaleb, A. K. Jain, "Face Detection in Color Images", *IEEE Computer Society*, Vol. 24(5), ISSN:0162-8828, pp. 696 - 706, 2002.

[21] J. B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations", *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, Vol. 1, pp. 281-297, 1967.

[22] M. T. Orchard and C. A. Bouman, "Color Quantization of Images", *IEEE Transactions on Signal Processing*, Vol. 39(12), pp. 2677-2690, 1991.

[23] Y. Y. Chuang, A. Agarwala, B. Curless, D. Salesin, R. Szeliski, "Video Matting of Complex Scenes", In *ACM SIGGRAPH 2004*, Vol. 21(3), pp. 243-248, 2002.



(a) (b) (c) (d) (e)

Figure 6: Bayes matting. (a) The original image; (b) The mask with hard edges of cutout; (c) The extracted foreground corresponding to (b); (d) The alpha mask with the soft edges of matting; (e) The matting result corresponding to (d).

Figure 7: Some other results of our algorithm. The upper row shows four original head shoulder pictures and the lower row presents the corresponding cutout results.

(a) (b) (c)

Figure 8: One cutout and composite example. (a) The original picture; (b) The cutout result; (c) The composite result.