Active Contour Image Segmentation in Fisher Discriminant Spaces

Ashoka Jayawardena
and Paul Kwan
School of Science and Technology
University of New England
Armidale, NSW, 2351
Email: {ashoka,kwan}@turing.une.edu.au

***Abstract*—In this paper, we introduce an algorithm that is able to segment objects in natural images by using active contours. Active contours are used to regularize the segmentations. Our approach utilizes multiple feature spaces to capture as much information as possible, followed by projecting the multiple dimensional features space onto a single dimension to enable improved active contour evolution. We apply the Fisher Linear Discriminant Analysis (FLDA) to optimally calculate the projection vector while providing prior knowledge on number of clusters that are present on the image. Preliminary experiments confirm that the proposed algorithm is able to segment objects in natural images while optimizing contour smoothness and noises.**

I. Introduction

Many scene analysis and processing applications require segmentation of a foreground object from its background. There are a number of popular approaches to achieve this task depending on the type of application. For example, general purpose image segmentation applications typically uses region growing [6] or active contour evolution [2], [4], [8], [12], [14] while background substraction [1], [7] is popular in foreground segmentation in video surveillance applications.

There are two advantages of using active contour methods for image segmentation. Firstly, active contours have been employed due to good regularization properties such as contour smoothness (or the length of the contour), distance etc. These regularization properties are useful to minimize segmentation noises. Secondly, some of the prior knowledge on the image can be encoded in the initial active contour.

The idea behind active contour image segmentation is that a contour evolves subject to constraints imposed by the image such as the image gradient. The active contour image segmentation algorithms can be implemented by using classical snakes [8], [14] or level sets [12]. In both of these implementations, active contours behave as energy minimizing curves and hence are formulated as energy minimization problems. When the curve (or the front) evolves in the normal direction of the curve, we arrive at the following evolution scheme [11]:

$$\frac{\partial \phi}{\partial t} + F|\nabla \phi| = 0 \tag{1}$$

The speed F is normally modelled by mean curvature, thus resulting in mean curvature motion. The mean curvature evolution equation is given by in [11] and [12] by:

$$\frac{\partial \phi}{\partial t} + |\nabla \phi| \mathtt{div}\left(\frac{\nabla \phi}{|\nabla \phi|}\right) = 0 \tag{2}$$

The mean curvature motion has been extensively used to model geometric flow. In the level sets implementation, the evolving curve is normally embedded in the zeroth level set. They have shown to be able to undergo automatic topological changes.

Active contour segmentation algorithms have been developed where the object mean can be used to discriminate textures in image processing [12].

II. Active Contour Segmentation

Let us denote $u : D_I \rightarrow R$ the image which we want segment to two partitions. Let us define the feature function $\mathbf{f} : D_I \rightarrow R^N$, where N is the number of feature dimensions, be the feature space in which the feature function is evaluated at each image location, i.e. pixels.

We will use the following external energy model:

$$\begin{aligned} &E_1(C, \mathbf{c}_1) + E_2(C, \mathbf{c}_2) = \\ &\int_{C_+} (\mathbf{f}(x,y) - \mathbf{c}_1)^T \mathbf{D} (\mathbf{f}(x,y) - \mathbf{c}_1) dxdy + \\ &\int_{C_-} (\mathbf{f}(x,y) - \mathbf{c}_2)^T \mathbf{D} (\mathbf{f}(x,y) - \mathbf{c}_2) dxdy \end{aligned} \tag{3}$$

where $\mathbf{c}_1$ and $\mathbf{c}_2$ are constants and $\mathbf{D}$ a diagonal matrix with positive values.

It can be verified that when $\mathbf{c}_1$ is the mean of the feature space $\mathbf{f}$ that is inside the contour C while $\mathbf{c}_2$ is the mean of the feature space $\mathbf{f}$ that is outside the contour C, $E(C, \mathbf{c}_1, \mathbf{c}_2) = E_1(C, \mathbf{c}_1) + E_2(C, \mathbf{c}_2)$ achieves its minimum for the given contour C. To see this, we calculate the partial derivatives of $E(C, \mathbf{c}_1, \mathbf{c}_2)$ with respect to $\mathbf{c}_1$ and $\mathbf{c}_2$:

$$\frac{\partial (E(C, \mathbf{c}_1, \mathbf{c}_2))}{\partial \mathbf{c}_1} = -2\mathbf{D} \int_{C_+} (\mathbf{f}(x,y) - \mathbf{c}_1) dxdy$$

$$\frac{\partial (E(C, \mathbf{c}_1, \mathbf{c}_2))}{\partial \mathbf{c}_2} = -2\mathbf{D} \int_{C_-} (\mathbf{f}(x,y) - \mathbf{c}_1) dxdy$$

When $\mathbf{c}_1$ and $\mathbf{c}_2$ are the mean of the inside and the outside regions of the contour C, the above two partial derivatives vanish. If we further assume that for each feature dimension the feature value of a texture object is approximately constant, it is clear that when C is the contour separating the objects $E(C, c_1, c_2)$ achieves its global minimum.

Let the step function H and the dirac impulse function δ be given by:

$$H(t) = \begin{cases} 1 & \texttt{if} \quad t \geq 0 \\ 0 & \texttt{if} \quad t < 0 \end{cases}$$

and

$$\delta(t) = \frac{d}{dt}(H(t))$$

respectively.

Then we can rewrite the external energy equation (3) as follows:

$$\begin{aligned} &E(C, \mathbf{c}_1, \mathbf{c}_2) = \\ &\int_{D_I} (\mathbf{f}(x,y) - \mathbf{c}_1)^T \mathbf{D} (\mathbf{f}(x,y) - \mathbf{c}_1) H(\phi(x,y)) dxdy + \\ &\int_{D_I} (\mathbf{f}(x,y) - \mathbf{c}_2)^T \mathbf{D} (\mathbf{f}(x,y) - \mathbf{c}_2)(1 - H(\phi(x,y))) dxdy \end{aligned} \tag{4}$$

In order to perform the gradient descent of the energy equation, we can decompose the evolution equation of $\phi_t(x, y)$ as:

$$\frac{\partial \phi}{\partial t} = \frac{\partial \phi}{\partial t}_{external} + \frac{\partial \phi}{\partial t}_{internal}$$

Now the external energy component of the evolution equation is given by:

$$\begin{aligned} &\frac{\partial \phi}{\partial t}_{external} \stackrel{def}{=} \frac{d(E(C))}{d\phi} = \\ &\delta(\phi) \left(-(\mathbf{f}(x,y) - \mathbf{c}_1)^T \mathbf{D} (\mathbf{f}(x,y) - \mathbf{c}_1) + \right. \\ &\left. (\mathbf{f}(x,y) - \mathbf{c}_2)^T \mathbf{D} (\mathbf{f}(x,y) - \mathbf{c}_2) \right) \end{aligned}$$

It is clear from the above equation that if the image domain consists only of a single texture object, then the external energy contribution to the evolution equation is zero. Therefore, only locally discriminative features contribute to the evolution equation.

The smoothness of the curve can be controlled by minimizing the length of the curve, thus yielding an external energy term for the evolution equation. Using the steps as outlined in [12], we arrive at the following level set formulation of the curve evolution:

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \delta_\epsilon(\phi) \left[\mu \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - v \right. \\ &- \lambda_1 (\mathbf{f}(x,y) - \mathbf{c}_1)^T \mathbf{D} (\mathbf{f}(x,y) - \mathbf{c}_1) \\ &\left. + \lambda_2 (\mathbf{f}(x,y) - \mathbf{c}_2)^T \mathbf{D} (\mathbf{f}(x,y) - \mathbf{c}_2) \right] \end{aligned} \tag{5}$$

A. Direction of External Energy Gradient

Given a pixel, (x, y), the sign of the external energy term determines whether that pixel is attracted to the positive side or the negative side of the level set function ϕ. Thus, if we choose two feature spaces where either of them can discriminate the partitions, each feature space can pull the pixel in different directions. This can lead to cancellation of the pull thus leading to unstable gradient descent or resulting in entirely arbitrary partitions.

The direction of grdient descent can be estimated by $sign(\mathbf{c}_1 - \mathbf{c}_2)$. Let $\mathbf{S}$ be a diagonal matrix with $sign(\mathbf{c}_1 - \mathbf{c}_2)$ as the diagonal. Thus, the evolution equation becomes

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \delta_\epsilon(\phi) \left[\mu \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - v \right. \\ &- \lambda_1 (\mathbf{f}(x,y) - \mathbf{c}_1)^T \mathbf{SD} (\mathbf{f}(x,y) - \mathbf{c}_1) \\ &\left. + \lambda_2 (\mathbf{f}(x,y) - \mathbf{c}_2)^T \mathbf{SD} (\mathbf{f}(x,y) - \mathbf{c}_2) \right] \end{aligned} \tag{6}$$

B. Segmentation Algorithm

We obtain the individual weights of the mutiple feature spaces by projecting the multidimensional feature space into a single dimension by using Fisher Linear Discriminant Analysis (FLDA). We restrict our attention to a two class linear discriminant analysis. The placement of the original contour provides us with a classificattion of pixels into either foreground or background. We can thus calculate the probability that a certain pixel value falls into the foreground or background. This enables us to roughly classify the pixels into the object that we hope to segment and its background. However, this could be improved by classifying the pixels into feature clusters, which is then followed by a classification of these feature clusters into the object that is to be segmented and the background. We can make use of a suitable clustering algorithm such as the k-means algorithm. The set of resulting clusters are further classified by referencing the initial contour seperating the foreground object and its background so that a projection vector can be computed using FLDA. Finally, the foreground object is segmented using an implicit active contour evolution algorithm that makes use of regularization features for the active contour evolution. The following summarises the resulting algorithm:

1) Classify the image into clusters $\{P_1, ..., P_K\}$.
2) Using the location of the initial contour, classify the clusters into foreground object $\{P_1^f, ..., P_{K_f}^f\}$ and the background object $\{P_1^b, ..., P_{K_b}^b\}$.
3) Project the original feature space by calculating the projection vector using $\{P_1^f, ..., P_{K_f}^f\}$ as the foreground class and $\{P_1^b, ..., P_{K_b}^b\}$ as the background class using FLDA.
4) Perform the active contour evolution untill convergence.

III. Linear Discriminant Spaces

The primary source that contributes to the construction of the weighting matrix $\mathbf{D}$, is the discrimination energy. The discrimination energy can be assessed by evaluating a certain

criterion function on the feature space. Unser [13] proposed the following criterion for optimizing texture transforms:

$$J_U = \frac{(\mu_1 - \mu_2)^2}{\mu_1 \mu_2}$$

where μ_1 and μ_2 are means of the objects to be segmented for a particular feature dimension.

Mahalanobis and Singh [9] suggested $J_{MS} = \frac{\mu_1}{\mu_2}$ which was shown to be not powerful enough for texture discrimination. On the other hand, a popular criterion used by the pattern recognition community is the Fisher's criterion given by

$$J_F = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}$$

where μ_1 and μ_2 are means of the objects to be segmented as before, and σ_1^2 and σ_2^2 are the variances of the objects for a particular feature dimension.

Here, We construct a variant of the fisher's criterion in order to construct the weighting matrix. The weighting matrix has the role of projecting the n-dimentional feature space onto a single dimensional feature space. That is, the diagonal of the weighting matrix is ω^2 of some projection vector ω.

The foreground and background are segmented into clusters. Here, let S_{b1} be the between class scatter matrix of the means of the clusters belonging to the foreground and S_{b2} the between class scatter matrix of the means of the clusters belonging to the background. Further, let S_w be the within class scatter matrix of all the clusters and μ_1 and μ_2 the means of the foreground and background respectively. We aim to maximize the following functional:

$$J(\omega) = \frac{\omega^T(\mu_1 - \mu_2)}{\alpha\omega^T(S_{b1} + S_{b2})\omega + \beta\omega^T S_w \omega}$$

Since we can scale the projection vector ω arbitrarily, we can equivalently write

$$\begin{aligned} &\arg\max_{\omega} \omega^T(\mu_1 - \mu_2) \\ &s.t. \\ &\alpha\omega^T(S_{b1} + S_{b2})\omega + \beta\omega^T S_w \omega = 1 \end{aligned}$$

Now the Lagrangian can be written as,

$$\omega^T(\mu_1 - \mu_2) - \lambda(\omega^T(\alpha(S_{b1} + S_{b2}) + \beta S_w)\omega - 1) = 0$$

Since solutions to the optimization problem occurs at the stationary points of the Lagrangian, we have

$$(\alpha(S_{b1} + S_{b2}) + \beta S_w)\omega = \mu_1 - \mu_2,$$

which yield

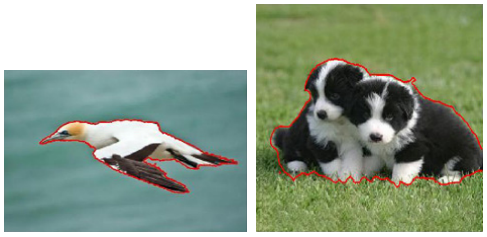
$$\omega = (\alpha(S_{b1} + S_{b2}) + \beta S_w)^{-1}(\mu_1 - \mu_2) \tag{7}$$

Many applications has this result in the form of $S_w^{-1}S_b$ where S_b is the between class scatter matrix and S_w is as before. These applications such as face recognition has very high dimensionality of the feature space and lack of sufficient samples to obtain full rank S_w as in [3], [5], [10]. However, our application involving image segmentation most likely has full rank of S_w.

The term βS_w in the equation (7) has two roles:

1) Regularize if $S_{b1} + S_{b2}$ is singular.
2) Minimize within cluster scatter.

The regularization feature is critical if we have less than three clusters for each of the foreground or background objects. However, the minimization of within cluster scatter is also important for discrimination.



(a) (b)

Fig. 1. (a) Segmented gannet image (b) Segmented puppies image.

IV. Results and Discussion

We performed experiments with two types of feature spaces, the original RBG space and a normalized colour space. Each of these spaces has advantages in natural image analysis. The RGB space was selected due to its superior performance relative to the normalized colour space when camouflage is present in the images, while normalized colour space was selected due to its superior performance when nonlinear lighting effects are present.

A. RGB Colour Space

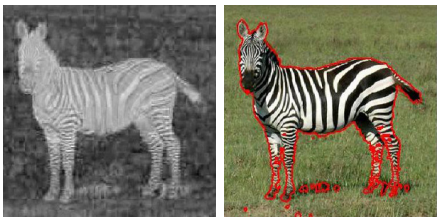
The results of the RGB colour space experiments on the gannet and puppies images are given in figure 1. The puppies image was segmented into three clusters while the gannet image was segmented into six clusters. The segmentation of both puppies and the gannet images was done without difficulties in terms of noise. The zebra image was clustered into five clusters due to substantial mixing of the background colour in the leg areas. Figure 2 illustrates the basic k-means clusters, the projected image, and active contour segmentation. In the segmentation, the leg areas of the zebra image still contain significant noises due to poor fit of the projection. The projection direction could be recalculated during the evolution which could result in a better fit of the projection direction. Figure 3 illustrates the improvements on the errorneous classification that could be obtained by recalculating the projection vector every five iterations of a 20 iterations run of the zebra image.

B. Normalized Colour Space

Normalized colour space has the ability to minimize shadowing affects that are present in coloured images of natural scenes.

(a) (b) (c)

Fig. 2. (a) K-means clusters. (b) Projected image. (c) Segmented image



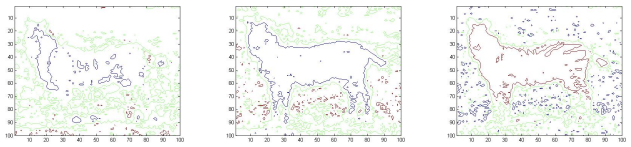
(a) (b)

Fig. 3. Recalculating the projection vector during the evolution. (a) Final Projected image. (c) Segmented zebra image

Here, we use a regularized colour space given by

$$\left(\frac{r+c}{r+g+b+3c}, \frac{g+c}{r+g+b+3c}, \frac{b+c}{r+g+b+3c}\right)$$

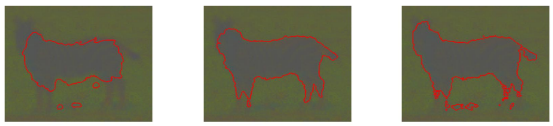
where (r, g, b) are red, green, and blue colour values of an image pixel and c some regularization constant. This regularized normalized colour space behaves smoothly when colour values are closer to zero. An example of such space is given in Figure 5. Figure 4 demonstrates the presence of different information in different dimension of the normalized colour space.



(a) (b) (c)

Fig. 4. (a) Blue colour channel. (b) Green colour channel. (c) Red colour channel

Fig. 5. Normalized colour Zebra image.



(a) (b) (c)

Fig. 6. Active Contour Segmentation in individual normalized colour dimensions. (a) Blue colour channel. (b) Green colour channel. (c) Red colour channel

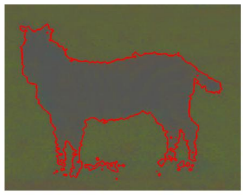


Fig. 7. Active Contour Segmentation in Fisher-space

We performed active contour segmentation in individual normalized colour dimensions (Figure 6) and in a Fisher's space (Figure 7) of the normalized colour space. The Fisher linear decomposition was performed using representative rectangular image patches obtained from two classes, i.e. foreground and background.

However, the projection approach we have described might have certain limitations. When all the cluster means (both foreground and background) are approximately aligned (i.e. located in a straight line) the projection will fail qualitatively due to the lack of a suitable direction to bring the cluster means of foreground or background closer.

V. Conclusions and Further Research

We have presented an algorithm that can be used for image segmentation when the objects of interest are composite, i.e. consist of multiple clusters. We have proposed a solution using the ideas of Fisher Linear Discriminant Analysis. Our algorithm successfully segment foreground objects when suitable projection directions exist. However, these limitations can be overcome by generalizing our active contour segmentation to multi-phase setting. Further research work, which we currently work on, includes unsupervised segmentation of foreground objects against background objects. The algorithm proposed in this paper can be used unsupervised, i.e. using a cluster number that is unrelated to the image at hand but sufficiently large, for simpler images.

References

[1] Elgammal A., Duraiswami R., Harwood D., and Davis L.S. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. In *Proceedings of IEEE*, number 7, pages 1151–1163, July 2002.

[2] Myronenko A. and Song X. Global active contour-based image segmentation via probability alignment. In *Computer Vision and Pattern Recognition, (CVPR'09)*, pages 2798–2804, 2009.

[3] Martinez A.M. Recognition of partially occluded and/or imprecisely localized faces using a probabilistic approach. In *Proceedings of Computer Vision and Pattern Recognition*, pages 712–717, Hilton Head, June 2000.

[4] Li C., Xu C., Gui C., and Fox M. D. Distance regularized level set evolution and its application to image segmentation. *IEEE Transactions on Image Processing*, 19(12):3243–3254, 2010.

[5] Swets D.D. and Weng J.J. Using discriminant eigenfeatures for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):831–836, 1996.

[6] Shapiro Linda G. and Stockman George C. *Computer Vision*. Prentice-Hall, 2001.

[7] Grimson W. E. L. and Stauffer C. Adaptive background mixture models for real-time tracking. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 22–29, 1999.

[8] Kass M., Witkin A., and Terzopoulos D. Snakes: Active contour models. *International Journal of Computer Vision*, 1:321–331, 1988.

[9] A. Mahalanobis and H. Singh. Application of correlation filters for texture recognition. *Applied Optics*, 33(11):2173–2179, 1994.

[10] Aleix M. Martinez and Avinash C. Kak. Pca versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):228–233, 2001.

[11] Osher S. and Sethian J. Front propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulation. *Journal of Computational Physics*, 79:12–49, 1988.

[12] Chan T.F. and Vese L.A. Active contours without edges. *IEEE Transactions on Image Processing*, 10:266–277, 2001.

[13] M. Unser. Local linear transforms for texture measurements. *Signal Processing*, 11(1):61–79, 1986.

[14] Caselles V., Catte F., Coll T., and Dibos F. A geometric model for active contour models in image processing. *Numerical Mathamatics*, 66:1–31, 1993.