

基於核轉換距離測量之膚色切割方法

鄭文昌、林宗毅、張家銘
修平技術學院資訊網路技術系
E-mail: wccheng@mail.hit.edu.tw

摘要

在此論文中我們提出一個新的影像膚色切割的方法，此方法主要是利用單一個高斯函數分類器來完成膚色切割，利用 Mahalanobis 距離做為輸入影像像素顏色與實際膚色的近似程度計算用來判斷輸入像素顏色是否為膚色，為了改善膚色的非線性分佈，我們利用核轉換推導出具有核轉換的單一高斯分類器的膚色切割方法，由於核轉換的方法已成功應用在許多問題並且有效改善問題的性能，因此在此論文我們結合色彩向量的核轉換方法完成影像中膚色的切割。從實驗結果發現，我們所提出的膚色切割方法能有效提升膚色切割的準確度。

關鍵詞：高斯分類器，Mahalanobis 距離、人臉偵測、色彩空間轉換。

一、簡介

膚色切割(Skin Segmentation)技術在電腦視覺及圖形辨識領域中扮演著重要的角色，例如人臉偵測、人臉追蹤、及手勢辨識都是常見的應用[1-4]。到目前為止已經有許多膚色切割技術被提出，而這些技術的大部分都是使用一個分類器將影像的像素依據顏色分成膚色(Skin)與非膚色(Non-skin)兩大群，因此隱藏在這些方法背後有一些前提假設及值得注意的議題，首先這些方法假設人類的膚色是有別於一般其它物體的顏色且膚色有相當高程度的一致性，因此進行膚色的切割對於上述這些應用才有意義，其次是不同色彩空間及不同分類器有不同的效果，到目前為止不同的色彩空間已經有被使用來做為膚色切割依據，在某些研究中提出色彩的亮度對於膚色切割並不相關[5]，因此只有使用顏色成分的資訊做為膚色的分類，總之，從以往的研究中我們發現四種色彩空間表示法最常被使用於膚色的切割，分別是 RGB、HSV、YCbCr、以及 CIE-Lab 色彩

空間，而其中又以 YCbCr 最常被使用[5]，因此在此論文中我們也採用 YCbCr 色彩空間做為膚色切割的依據。

另一方面，對於膚色的分類也已經有許多方法提出，它們包含有片段直線分類器(Piecewise Linear Classifiers)[6-9]、基於色彩直狀圖之貝氏分類器(Bayesian Classifier with Histogram)[2, 10]、高斯函數分類器(Gaussian Classifiers)[11-13]，以及類神經網路[14]，這些分類器所決定的邊界可以從很簡單的正方形邊界、球形邊界、橢圓形邊界甚至到更複雜邊界。對於片段直線分類器，這種分類器利用片段線性方程式邊界區分膚色與非膚色兩群，雖然計算簡單且容易實現，但是對於落在片段直線邊界膚色區域內的每一個值均有相同權重，換句話說，對於膚色的機率分佈都視為相同，因此無法考慮膚色的機率分佈，這是不合理的，改善的方法則是利用基於色彩直狀圖之貝氏分類器或者高斯分類器，這兩類的方法都以高斯函數為基礎的分類器，最大不同則在於當我們事先訓練貝氏分類器時，我們需要事先收集膚色與非膚色兩群的顏色資料，而高斯分類器則只需要膚色資料，由於膚色容易收集具代表性的資訊，而非膚色資料要收集具代表性相較於膚色資料較不容易。對於類神經網路分類器的訓練也有與貝氏分類器有相同問題。

在此論文我們採用單一高斯分類器做為膚色切割的分類器。雖然從以前研究知道單一高斯函數的分類器能有效的完成膚色切割，但是由於高斯函數的指數部分的 Mahalanobis 距離是屬於線性的距離測量，因此對於線性的膚色分佈可以有效且高準確的分類，但是如果膚色的分佈為非線性時，將會影響膚色切割結果，因此解決的方法是我們將輸入的色彩空間向量先透過非線性的核(Kernel)轉換[15]，將色彩空間向量透過轉換到另一個更高階的特徵向量空間，以期在此特徵空間有線性表現，

由於核轉換已成功應用在許多訊號處理的問題上，因此在此研究中我們提出結合核轉換及單一高斯函數分類器的膚色切割方法。

最後文章剩下部分的結構如下，第二章將介紹我們提出的膚色切割方法，第三章為實驗的部分結果，最後章節為結論。接下來我們將一一介紹。

二、膚色切割的方法

當影像輸入後，將每一個像素的 RGB 色彩空間向量值轉換成 YCbCr 色彩空間向量值，其轉換方程式可表示如下：

$$Y = \frac{R + G + B}{3},$$

$$\begin{bmatrix} Cb \\ Cr \end{bmatrix} = \begin{bmatrix} -0.1687 & -0.3313 & 0.5 \\ 0.5 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 128 \\ 128 \end{bmatrix}, \quad (1)$$

為了判別輸入影像中每一像素的 YCbCr 值是屬於膚色或非膚色類群，我們必須事先訓練具核轉換的高斯函數分類器參數。首先從影像中蒐集一組已知的膚色 YCbCr 的樣本點 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_M]$ ， $\mathbf{x}_i \in R^3$ ，令 \mathbf{z} 代表輸入影像像素在 YCbCr 色彩空間中的任意點，則膚色的切割可以利用計算 \mathbf{z} 與 \mathbf{m} 的指數的 Mahalanobis 距離來決定近似的程度，單一高斯函數分類器可以定義如下：

$$d(\mathbf{z}, \mathbf{m}) = \exp\left[-\frac{1}{2}(\mathbf{z} - \mathbf{m})^T \mathbf{C}^{-1}(\mathbf{z} - \mathbf{m})\right], \quad (2)$$

其中 $\mathbf{m} = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i$ ， $\mathbf{m} \in R^n$ 為膚色樣本點的平均向量 (Mean Vector)， \mathbf{C} 是膚色樣本點 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]$ 的共變異數矩陣 (Covariance Matrix)。因此對於任意像素 $\mathbf{z} = [Y, Cb, Cr]^T$ 向量值輸入至方程式(2)，並且依據方程式(3)決定是否為膚色結果：

$$Skin(\mathbf{z}) = \begin{cases} 1 & , d(\mathbf{z}, \mathbf{m}) > d_0 \\ 0 & , otherwise \end{cases} \quad (3)$$

當 $d(\mathbf{z}, \mathbf{m})$ 的值小於指定的臨界值 d_0 時，則輸入像素 \mathbf{z} 為膚色，反之為非膚色。

如同前述，雖然單一高斯函數的分類器能有效完成影像像素的膚色分類，但是高斯函數的指數部分的 Mahalanobis 距離是一線性距離測量，因此對於線性的膚色分佈可以有效且高準確的分類，但是如果膚色的分佈為非線性時，膚色切割準確性將會有所下降，因此我們利用一個非線性函數將輸入色彩空間向量轉換成新的特徵向量，使得在新的特徵空間中的得到線性的距離測量，架設存在一個非線性函數 $\Phi: \mathbf{x} \in R^n \rightarrow \bar{\mathbf{x}} \in R^N \subseteq F$ ，則單一高斯函數分類器可以寫成：

$$D(\Phi(\mathbf{z}), \boldsymbol{\mu}) = \exp\left\{-\frac{1}{2}(\Phi(\mathbf{z}) - \boldsymbol{\mu})^T \tilde{\mathbf{C}}^{-1}(\Phi(\mathbf{z}) - \boldsymbol{\mu})\right\}$$

$$= \exp\left\{-\frac{1}{2}\tilde{\Phi}(\mathbf{z})^T \tilde{\mathbf{C}}^{-1}\tilde{\Phi}(\mathbf{z})\right\}, \quad (4)$$

其中 $\tilde{\Phi}(\mathbf{z}) = \Phi(\mathbf{z}) - \boldsymbol{\mu}$ ， $\boldsymbol{\mu} = \sum_{i=1}^M \Phi(\mathbf{x}_i) / M$ 是在 F 空間的膚色特徵向量的平均向量， $\tilde{\mathbf{C}} = \sum_{i=1}^M \tilde{\Phi}(\mathbf{x}_i) \tilde{\Phi}(\mathbf{x}_i)^T / M$ 則是在 F 空間的膚色特徵向量的共變異矩陣，因此我們可以找到共變異矩陣的特徵向量 $\lambda \geq 0$ 及特徵向量 $\mathbf{v} \in F$ 使得滿足 $\lambda_j \mathbf{v}_j = \tilde{\mathbf{C}} \mathbf{v}_j$ ， $j = 1, 2, \dots, N$ 。由於共變異矩陣的特徵向量 \mathbf{v}_j ， $j = 1, 2, \dots, N$ ，可以視為是落在由 $\tilde{\Phi}(\mathbf{x}_1), \tilde{\Phi}(\mathbf{x}_2), \dots, \tilde{\Phi}(\mathbf{x}_M)$ ，為基底所展開的向量空間，因此存在對應於每個基底的係數 $\alpha_1^j, \alpha_2^j, \dots, \alpha_i^j, \dots, \alpha_M^j$ 使得可以將任一共變異矩陣的特徵向量表示成基底的線性組合 $\mathbf{v}_j = \sum_{i=1}^M \alpha_i^j \tilde{\Phi}(\mathbf{x}_i)$ 。而這些對應於每一個基底 $\tilde{\Phi}(\mathbf{x}_1), \tilde{\Phi}(\mathbf{x}_2), \dots, \tilde{\Phi}(\mathbf{x}_M)$ 的係數所形成的向量可以利用計算 \mathbf{K} 矩陣的特徵向量而求得，在 [k-96] 中將這些係數表示成下式

$$\mathbf{P} = [\boldsymbol{\alpha}^1, \boldsymbol{\alpha}^2, \dots, \boldsymbol{\alpha}^N] \in R^{M \times N}. \quad (5)$$

且滿足 $\lambda_j \boldsymbol{\alpha}^j = \mathbf{K} \boldsymbol{\alpha}^j$ ， $j = 1, 2, \dots, N$ 。我們重新整理這些方程式，將上述的推倒結果帶回方程式(4)中，我們可以得到在新的特徵空間 F 的單一高斯函數分類器如下：

$$\begin{aligned}
D(\Phi(\mathbf{z}), \boldsymbol{\mu}) &= \exp\left\{\tilde{\Phi}(\mathbf{z})^T (\mathbf{V}\mathbf{D}^{-1}\mathbf{V}^T)\tilde{\Phi}(\mathbf{z})\right\} \\
&= \exp\left\{(\mathbf{V}^T\tilde{\Phi}(\mathbf{z}))^T \mathbf{D}^{-1}(\mathbf{V}^T\tilde{\Phi}(\mathbf{z}))\right\} \\
&= \exp\left\{\left[\sum_i \alpha_i^1 \tilde{K}(\mathbf{x}_i, \mathbf{z}) \quad \cdots \quad \sum_i \alpha_i^N \tilde{K}(\mathbf{x}_i, \mathbf{z})\right] \right. \\
&\quad \left. \cdot \mathbf{D}^{-1} \begin{bmatrix} \sum_i \alpha_i^1 \tilde{K}(\mathbf{x}_i, \mathbf{z}) \\ \sum_i \alpha_i^2 \tilde{K}(\mathbf{x}_i, \mathbf{z}) \\ \vdots \\ \sum_i \alpha_i^N \tilde{K}(\mathbf{x}_i, \mathbf{z}) \end{bmatrix}\right\},
\end{aligned} \tag{6}$$

其中

$$\begin{aligned}
\tilde{K}(\mathbf{y}, \mathbf{z}) &= \tilde{\Phi}(\mathbf{y})^T \tilde{\Phi}(\mathbf{z}) \\
&= \left(\Phi(\mathbf{y}) - \frac{1}{M} \sum_{i=1}^M \Phi(\mathbf{x}_i)\right)^T \left(\Phi(\mathbf{z}) - \frac{1}{M} \sum_{k=1}^M \Phi(\mathbf{x}_k)\right) \\
&= K(\mathbf{y}, \mathbf{z}) - \frac{1}{M} \sum_{j=1}^M K(\mathbf{x}_j, \mathbf{z}) - \frac{1}{M} \sum_{k=1}^M K(\mathbf{y}, \mathbf{x}_k) \\
&\quad + \frac{1}{M^2} \sum_{j,k=1}^M K(\mathbf{x}_j, \mathbf{x}_k)
\end{aligned} \tag{7}$$

因此對於影像中任意一像素 $\mathbf{z} = [Y, Cb, Cr]^T$ 經過非線性轉換得 $\Phi(\mathbf{z})$ ，再利用方程式(6)計算在新的特徵向量空間中 $\Phi(\mathbf{z})$ 與膚色的相似程度，當距離小於指定的臨界值 D_0 時，則 $\Phi(\mathbf{z})$ 在 F 空間是一個膚色。

三、實驗結果

為了訓練單一高斯分類器的參數，首先我們從以往拍攝的影像資料庫及網路收集的影像資料庫中以人工方式任意擷取膚色共 750 筆資料，利用此膚色資料計算 $K(\mathbf{x}_i, \mathbf{x}_j)$ 在核轉換後的特徵空間的特徵值與特徵向量，帶回方程式(6)中準備用來計算任意輸入的彩色影像中每一個像素的 $\Phi(\mathbf{z})$ 與膚色的在轉換後特徵空間中的近似程度，從上一章節中我們知道不同臨界值將決定最後膚色切割的結果，因此第一個實驗我們將透過實驗決定是當臨界值以利用在後續的實驗中。圖 1 為一個膚色切割範例，圖 1(a)為原始彩色影像，圖 1(b)-(e)為膚色偵測結果，結果圖中的亮點即為膚色，從圖 1 得到當臨界值設為 0.90~0.92 之間，有較好的膚色切割結果。

此外為了證明我們提出的膚色切割方法的性能，我們將我們提出方法的結果與片段直線分類器及單一高斯分類器比較，圖 2 呈現為部分實驗影像的膚色切割結果，圖 2 中(a)欄位為原始影像，圖 2(b)欄位為片段直線分類器的膚色切割結果，圖 2(c)欄位為單一高斯分類器膚色切割結果，圖 2(d)欄位為我們提出具有核轉換單一高斯分類器的膚色切割結果，從圖 2 中我們發現，片段直線分類器的結果最差，而單一高斯分類器與我們提出改良式單一高斯分類器結果近似，但是以膚色切割完整性來說，我們的方法比單一高斯分類器更好，例如圖 2 中第二列的例子，我們方法完整切割出影像中人像右上手臂內側，而單一高斯分類器無法完整切割，又例如圖 2 中第六個例子中，人臉膚色切割結果我們提出方法比起單一高斯分類器有較完整的切割結果。



(a)



(b)



(c)



(d)



(e)

圖 1. 膚色偵測的範例，(a) 原始彩色影像，膚色偵測結果，臨界值為 (b) 0.88，(c) 0.90，(d) 0.92，(e) 0.94。

四、結論

在此論文中我們依據單一高斯函數分類器提出一個具有核轉換距離測量的高斯函數分類器來完成影像中膚色的切割，將輸入色彩空間向量透過非線性的核(Kernel)轉換將 YCbCr 色彩空間向量轉換到另一個更高階的特徵向量空間，使得在此特徵空間中能有線性表現的距離測量。從實驗結果發現，具有核轉換的單一高斯函數分類器能有效提升膚色切割的準確度。

對於未來的研究將包含兩個部分，一為持續改善膚色切割方法的性能，利用具有核轉換的多高斯函數分類器考慮多種人種膚色的情況，另一方面將會將此膚色切割方法應用在人臉偵測、人臉辨識、以及手勢辨識應用上，以提升智慧型環境安全監控系統的性能。

致謝

此研究結果由國科會計畫所支助，計畫名稱為區域空間之智慧型整合式安全監控系統，編號為 NSC 96-2221-E-164-011-。

五、參考文獻

- [1] E. Hjelm and B. K. Low, "Face Detection: A Survey," *Computer Vision and Image Understanding*, vol. 83, no. 3, pp. 236-274, Sept. 2001.
- [2] M.-H. Yang, D. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34-58, Jan. 2001.
- [3] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proc. IEEE*, vol. 83, pp. 705-740, May. 1995.
- [4] W. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, "Face Recognition: A Literature Survey," CVL Technical Report, Center for Automation Research, Univ. of Maryland at College Park, Oct. 2000.
- [5] S.-L. Phung, A. Bouzerdoum, and D. Chai, "Skin Segmentation Using Color Pixel Classification: Analysis and Comparison," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, Jan. 2005.
- [6] D. Chai and K. N. Ngan, "Face Segmentation Using Skin Color Map in Videophone Applications," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 9, no. 4, pp. 551-564, 1999.
- [7] C. Garcia and G. Tziritas, "Face Detection Using Quantized Skin Color Region Merging and Wavelet Packet Analysis," *IEEE Trans. on Multimedia*, vol. 1, no. 3, pp. 264-277, 1999.
- [8] K. Sobottka and I. Pitas, "A Novel Method for Automatic Face Segmentation, Facial Feature Extraction and Tracking," *Signal Processing: Image Comm.*, vol. 12, no. 3, pp. 263-281, 1998.
- [9] R.-L. Hsu, A.-M. Mohamed and A. K. Jain, "Face Detection in Color Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696-706, May. 2002.
- [10] H. Wang and S. F. Chang, "A Highly Efficient System for Automatic Face Detection in Mpeg Video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 7, no. 4, pp. 615-628, 1999.
- [11] H. Greenspan, J. Goldberger, and I. Eshet, "Mixture Model for Face Color Modeling and Segmentation," *Pattern Recognition Letters*, vol. 22, pp. 1525-1536, Step. 2001.
- [12] B. Menser and M. Wien, "Segmentation and Tracking of Facial Regions in Color Image Sequences," *SPIE Visual Comm. and Image Processing*, vol. 4067, pp. 731-740, June 2000.
- [13] M.-H. Yang and N. Ahuja, "Gaussian Mixture Model for Human Skin Color and Its Application in Image and Video Databases," *SPIE Storage and Retrieval for Image and Video Databases*, vol. 3656, pp. 450-466, Jan. 1999.
- [14] S.L. Phung, D. Chai, and A. Bouzerdoum, "A Universal and Robust Human Skin Color Model Using Neural Networks," *Proc. of INNS-IEEE Internal Joint Conf. Neural Networks*, vol. 4, pp. 2844-2849, July 2001.
- [15] András Kocsor and László Tóth, "Kernel-Based Feature Extraction with a Speech," *IEEE Trans. on Signal Processing*, Vol. 52, No. 8, pp. 2250-2263, August 2004.



圖 2. 部分測試影像的膚色偵測結果的比較，(a)原始彩色影像，膚色切割結果藉由(b)片段線性方法[5] (c)單一高斯函數方法 (d)具核轉換距離的單一高斯函數方法。