Engineering Applications of Artificial Intelligence 32 (2014) 136–150

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Image skin segmentation based on multi-agent learning Bayesian and neural network

A.A. Zaidan [a,*], N.N. Ahmad [a], H. Abdul Karim [a], M. Larbani [b], B.B. Zaidan [a], A. Sali [c]

[a] Faculty of Engineering, Multimedia University, Cyberjaya, Selangor Darul Ehsan, Malaysia
[b] Department of Business Administration, IIUM University, Jalan Gombak, Kuala Lumpur, Malaysia
[c] Department of Computer and Communication Systems, Faculty of Engineering, University Putra Malaysia, Serdang, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 9 January 2013
Received in revised form
17 October 2013
Accepted 6 March 2014
Available online 5 April 2014

Keywords:
Skin detector
Bayesian method
Neural network

ABSTRACT

Skin colour is considered to be a useful and discriminating spatial feature for many skin detection-related applications, but it is not sufficiently robust to address complex image environments because of light-changing conditions, skin-like colours and reflective glass or water. These factors can create major difficulties in face pixel-based skin detectors when the colour feature is used. Thus, this paper proposes a multi-agent learning method that combines the Bayesian method with a grouping histogram (GH) technique and the back-propagation neural network with a segment adjacent-nested (SAN) technique based on the YCbCr and RGB colour spaces, respectively, to improve skin detection performance. The findings from this study have shown that the proposed multi-agent learning for skin detector has produced significant true positive (TP) and true negative (TN) average rates (i.e. 98.44% and 99.86% respectively). In addition, it has achieved a significantly lower average rate for the false negative (FN) and false positive (FP) (i.e. only 1.56% and 0.14% respectively). The experimental results show that multi-agent learning in the skin detector is more efficient than other approaches.

1. Introduction

Skin detection is one of the most important techniques in image processing and is the most distinctive and widely used approach for many applications (Wadud et al., 2008; Mohamed et al., 2008), such as face detection (Dargham et al., 2009), human motion analysis (Peer et al., 2009), pornographic image filters (Daxiang et al., 2013), and blocking objectionable content (Ottinger et al., 2009). Correspondingly, applications such as content-aware video compression and image colour balancing can also benefit from automatic skin detection in images. Skin detection is used to determine the image pixels that are related to human skin. Colour is a useful cue for extracting skin pixels, and choosing a skin colour as an element in detecting human presence is a relatively simple and straightforward task; in addition, skin colour has a processing time advantage, given that colour processing is faster compared to other processing of other features (Taqa and Jalab, 2010b); nevertheless, using colour information for skin detection is challenging because the appearance of skin in images is affected by various factors, such as the illumination, background, camera characteristics, and reflections from glass or water (Liu et al., 2011; Shoyaib et al., 2012; Khan et al., 2012; Mohair et al., 2012). Many techniques have been presented in the literature for skin detection using colour (Bhoyar and Kakde, 2010). Skin detection is the process of finding skin-coloured pixels and regions in an image or a video (Elgammal et al., 2009). This process is typically used as a pre-processing step to finding regions that potentially have human faces and limbs in the images. A skin colour detector typically transforms a given pixel into an appropriate colour space and then uses a skin classifier to label the pixel, whether it is a skin or a non-skin pixel (Berbar, 2011). A skin classifier defines a decision boundary of the skin colour class in the colour space based on a training database of skin-coloured pixels. The main goal of a skin colour detection algorithm is to build a decision rule that will discriminate between skin and non-skin pixels (Zolfaghari et al., 2011). This goal is accomplished by introducing a metric that measures the distance between pixel colours in relation to the skin tone. This metric type is defined by the skin colour modelling method. Different methods for discriminating between skin and non-skin pixels are available in the literature (Vezhnevets et al., 2003; Kakumanu et al., 2007). However, it was only in the last 15 years that the value of skin colour segmentation has been fully realised in various computer vision applications (Naji et al., 2012). These applications can be grouped into three types of skin modelling, namely, explicitly defined skin regions, non-parametric skin, and parametric skin approaches.

The first category uses explicit rules based on colour values. The explicit skin cluster classifier experimentally defines the

* Corresponding author.
E-mail address: aws.alaa@gmail.com (A.A. Zaidan).

http://dx.doi.org/10.1016/j.engappai.2014.03.002
0952-1976/

boundaries of the skin cluster in a certain colour space (Vezhnevets et al., 2003). These methods use explicit rules that are based on colour values, which is an approach that many studies have taken (Kovac et al., 2003). These methods are very modest to implement, and the simplicity of the skin detection rules leads to a quick and inexpensive construction of a classifier. However, the main difficulty of achieving high recognition rates with this method is the need to find empirically both good colour space and adequate decision rules. Recently, a method that uses machine learning (ML) algorithms to find a suitable colour space and a simple decision rule and thus achieves high recognition rates was proposed (Sisodia and Verma, 2011).

The second category uses non-parametric models. These modelling methods are better suited for constructing classifiers in the cases of limited training and expected target datasets (Shoyaib et al., 2012). The generalisation and interpolation ability of these methods facilitate the construction of classifiers that have acceptable performance using incomplete training data (Surendiran and Alagarsamy, 2010; Paul et al., 2011). These methods are expressed with a small number of parameters and need very little storage space; however, it requires more time to compute the skin probability model in comparison to non-parametric methods. Thus, they can be very slow; a mixture of Gaussians in both training and work, as well as in their performance, depends strongly on the skin distribution shape. Moreover, most of the parametric skin modelling methods ignores the non-skin colour statistics. This aspect, together with the dependence on the skin cluster shapes, results in higher false positive rates compared to non-parametric methods (Vezhnevets et al., 2003). Neural Networks are considered to be the most suitable method to use in parametric skin modelling compared with other methods (Bhoyar and Kakde, 2010; Doukim et al., 2010; Wylie, 2010; Zolfaghari et al. 2011).

The third category uses parametric models for skin colour distributions. The key idea of non-parametric skin modelling methods is to estimate skin colour distributions from the training data without deriving an explicit model of the skin colour. The results of these methods are sometimes referred to as the skin probability map, which assigns a probability value to each point in a discrete colour space (Brand and Mason, 2000; Gomez, 2002). Vezhnevets et al. (2003) and Kakumanu et al. (2007) demonstrated three clear advantages of non-parametric methods: fast training, classification, and usage. Their performance depends heavily on the training set selection and is theoretically independent of the shape of the skin distribution, which is not the case for explicit skin cluster definitions and parametric skin modelling. These methods usually have high true positive and low false positive rates, which indicates that they can find most of the skin regions when the number of non-skin pixels marked as skin pixels is low (Kakumanu et al., 2007). The disadvantages of these methods are that they need a very large amount of storage space and cannot interpolate or generalise the training data. The storage of the (skin|RGB) table, also known as the lookup table (LUT), requires a large amount of memory (Vezhnevets et al., 2003). A solution to this problem is to use a smaller colour space, such as a colour space with 32^3 colours instead of 256^3 colours (Phung et al., 2005). A Naïve Bayesian classifier, however, is considered to be the most suitable method to be used in non-parametric skin modelling compared with the others (Zhanyu and Leijon, 2010; Cao and Liu, 2011; Liu et al., 2011).

This paper proposes a multi-agent learning system that combines the Bayesian method with a grouping histogram technique and the back-propagation neural network with a segment adjacent-nested technique based on YCbCr and RGB colour spaces, to improve the skin detection performance.

The remaining sections are organised as follows. Section 2 covers the related research and objectives of the present study. The proposed system design is described in Section 3. In Section 4, the experiments conducted for the proposed skin detector are described. An evaluation and discussion of the results are provided in Section 5, followed by the research limitations, the research contributions, the future works and conclusions in Sections 6, 7, 8 and 9, respectively.

2. Literature review

A number of skin colour segmentation approaches have been proposed based on neural networks or Naïve Bayesian classifiers (Bhoyar and Kakde, 2010).

Duan et al. (2009) proposed a method of human skin region detection that is based on a pulse-coupled neural network (PCNN). The original input image was translated from RGB colour space to YIQ colour space, and then, a channelled image is obtained. Elgammal et al. (2009) used the Bayesian approach for skin detection based on an RGB colour space. Wylie (2010) used neural networks for skin detection to detect skin pixels based on an RGB colour space. Their system was, however, prone to picking up objects that are similar in colour, such as hair, which is skin-like. Wylie has asserted that there is a positive relationship between the number of perceptrons in the hidden layer and the performance of the skin detector. Bhoyar and Kakde (2010) presented a pixel-based skin colour classification approach for detecting skin pixels and non-skin pixels in colour images by using a novel neural network symmetric classifier based on an YCbCr colour space. Zhanyu and Leijon (2010) proposed a Bayesian estimation algorithm using parameters that are based on an RGB space as the features, and the posterior mean was used as the point estimate of the parameters. The proposed skin detectors are still affected by different lighting conditions. More reliable skin detectors that can work effectively under different lighting conditions are needed.

In Doukim et al. (2010), two types of combination strategies were evaluated, namely combining skin features and combining skin classifiers. A major issue in the design of a multi-layer perceptron neural network based on an YCbCr colour space was to determine the optimal number of hidden units given a set of training patterns. The skin detection performance based on their combination strategies are not accurate, as the false is 14.90% due to the mean square error (MSE) results are consistent for few number of neurons in the hidden layer and then the variation occurs until it gives the lowest MSE value of 0.0425 and this is mean that the training error still high due to overlap between the training data.

Taqa and Jalab (2010b) proposed three skin detectors that were based on pre-defined rules of skin colour tones, texture features and a combination of both colour and texture features. In their work, a back-propagation artificial neural network-based RGB colour space was then used to learn the features and to classify any given input, each image the matrices R, G and B are converted into a vector, one parallel after another. This leads to the large training file and it cannot take the number of pixels from different images with higher training probabilities. This also prevents the establishment of training a sufficient number of pixels for a more credible region in the detection of skin. This shows the limitations of this approach.

Cao and Liu (2011) presented a skin colour classification technique that employs a Bayesian decision approach with colour statistics data. The statistics of the skin colour distribution were obtained in YCbCr colour space, and more studies under different lighting conditions must be performed.

Liu et al. (2011) proposed a method for detecting skin colours under varying illumination using a local edge-based colour constancy algorithm based on RGB colour space. However,