Comparative Performance of Different Chrominance Spaces for Color Segmentation and Detection of Human Faces in Complex Scene Images

Jean-Christophe Terrillon and Shigeru Akamatsu
ATR Human Information Processing Laboratories
2-2 Hikaridai, Seika-cho,
Soraku-gun, Kyoto 619-0288, Japan
{terril, akamatsu}@hip.atr.co.jp

Abstract

Color is a powerful fundamental cue that can be used at an early stage to detect objects in complex scene images. This paper presents an analysis of the performance of nine different chrominance spaces in the specific problem of automatically detecting and locating human faces in two-dimensional still scene images. For each space, we use a skin color model based on the Mahalanobis metric to segment faces from the scene background by thresholding. We perform feature extraction on the segmented images by use of fully translation-, scale- and in-plane rotation-invariant moments that are derived from the Fourier-Mellin transform, and apply a multilayer perceptron neural network with the invariant moments as the input vector to distinguish faces from distractors. We show that for each chrominance space, the detection efficiency is critically dependent on the goodness of fit of the skin chrominance distribution to the proposed model, and to a lesser extent on the discriminability between skin and "non-skin" distributions. Also, normalized color spaces are shown to yield the best segmentation results, and subsequently the highest rate of detection of faces with a large variety of poses and against relatively complex backgrounds.

1. Introduction

Automatic detection and localisation of human faces in two-dimensional natural, complex scene images is a difficult task that has been relatively unexplored until recently [7] [6] [1]. Face detection has important applications, as a first step in higher-level face recognition tasks such as personal identification for security purposes, the determination of sex and race, the understanding of facial expressions, or in the field of multimedia, for portrait retrieval in a large database of images or for interactive human-machine interfaces. In recent years, an increasing body of research has addressed the specific problem of automatic face detection based on skin color [8] [18] [17] [2] [21]. Color is a powerful fundamental cue that can be used as a first step in the process of face detection in complex scene images because color image segmentation is computationally fast while being relatively robust to changes in illumination, in viewpoint, in scale, to shading and to complex (cluttered) backgrounds as compared to the segmentation of grey-level images. Robustness is achieved if a color space efficiently separating the chrominance from the luminance in the original color image and a plausible model of the chrominance distribution of human skin are used for thresholding. In general, dimensionality reduction is first achieved by a suitable (linear or nonlinear) transformation from a 3-D RGB color space into a 2-D chrominance space (and into a separate luminance component). Normalized r-g chrominance space has often been used for face detection [8] [17] [21] [4] [15] because it reduces the sensitivity of the segmentation to changes in illumination. Other chrominance-luminance spaces that have been commonly used are the perceptually plausible HSV (or HSI) space [16] [9] [14] or the hardware-oriented YIQ or YES spaces [20] [3] [13]. In [10], the comparative efficiency of three different color spaces (HSI, CIE-L*u*v* and Karhunen-Loeve) in discriminating between skin (for Asian subjects only) and different facial features (mouth, eyes and eyebrows) has been analyzed, after a face has been segmented from a background. However, to our knowledge, no analysis of the comparative efficiency of several different chrominance spaces has been performed until now in the general problem of face detection. In effect, the efficiency of the color segmentation of a human face depends on the chrominance space that is selected, because the skin chrominance distribution depends on the chrominance space. Therefore, the selection of an appropriate color space is an important task.

We propose to compare the efficiency of nine different color spaces for face segmentation and detection against complex backgrounds. In section 2, we examine the chrominance