Abstract

Face recognition (FR) has been one of the most active research topics in computer vision for more than three decades. It has been widely applied in various practical scenarios such as access control system, massive surveillance, human computer interaction, etc. A conventional FR system consists of four stages, face detection, face alignment, face representation and face matching. Compared with the other three, face representation significantly affects the performance of a FR system because it determines whether the FR system is robust to real world variations, such as illumination, pose, occlusion, etc. Recently, research efforts have been dedicated to incorporate color information into a multiple-feature encoding scheme for improving FR performance. Specifically, multiple global and local feature representations are extracted from face images in a certain color space and then fused together for classification. Major challenges of such color FR tasks include how to construct an effective color space to represent color images, and how to fuse different feature representations extracted from face images. To tackle the challenge of color space construction, we propose a framework to derive an effective color space LuC_1C_2 from RGB. As for feature fusion of color features, we propose a Covariance Matrix Regularization (CMR) method, a Color Channel Fusion (CCF) method and a color face descriptor, Ternary Color Local Binary Patterns (TCLBP). More recently, Convolutional Neural Networks (CNNs) have been proved effective for extracting high-level visual features from face images. However, there exist significant problems in CNN feature representations. For example, the generalization ability of CNN features is limited when training and testing data has large differences, the feature fusion of different CNN architectures has not been thor-

oughly studied yet, and the performance of CNNs drops dramatically under the condition of low-resolution face images. To enhance the generalization ability of CNN features, we propose to combine CNN representations with color pixel values by score fusion. For different CNN architectures, we propose a simplified ResNet model, ResNetShort, to fuse its features with those of VGG-Face by CMR. For low-resolution FR (LRFR), we propose a Deep Coupled ResNet model (DCR).

Color in the machine vision system is defined by a combination of 3 color components specified by a color space. Existing color spaces are based on different criteria and their performance is not consistent on different datasets. This motivates us to propose a framework for constructing effective color spaces. The proposed color space, LuC_1C_2, consists of one luminance component and two chrominance components. The luminance component Lu is selected among four luminance candidates from existing color models by analysing their R,G,B coefficients and the color sensor properties. The chrominance components are derived by the discriminant analysis and the covariance analysis. Experiments show that both hand-crafted and CNN feature representations extracted from the LuC_1C_2 images perform consistently better than those extracted from images of other color spaces.

The fusion of multiple color features is important for achieving state-of-the-art FR performance. Existing color feature fusion methods either reduce the dimensionality of each feature first and then concatenate all low-dimensional feature vectors, named as DR-Cat, or the vice versa, named as Cat-DR. In DR-Cat, existing methods simply reduce features in different color channels to the same number of dimensions and concatenate them. But the importance or reliability of features in different color channels is not the same. We propose a Color Channel Fusion (CCF) approach to select more features from reliable and discriminative channels. DR-Cat ignores the correlation information between different features while Cat-DR fully uses it. This correlation information estimated from the training data may not be reliable. We propose a Covariance Matrix Regularization (CMR) technique to regularize the feature correlation estimated from training data before using it to train the feature fusion model. To better utilize the

correlation information between different color channels at the feature level, we propose the Ternay Color LBP (TCLBP) descriptor. Other than intra-channel LBP features, we extract the inter-channel LBP features by encoding the spectral structure of R,G,B component images at the same location.

CNNs have very large numbers of parameters that must be learned from millions of training examples. For every different application scenario, the common practice is to train a CNN on a very large dataset and then use the CNN either as a fixed feature extractor or an initialization for fine-tuning on images from the application of interest. The non-CNN features, color pixel values, depict the characteristics of face images from a different perspective. Thus color pixel values provide information complementary to the image representations learned by CNNs. We propose to use color pixel values together with fine-tuning to boost the generalization ability of pre-trained CNNs. To further improve the performance of CNNs, we train a simplified ResNet model, ResNetShort, and combine its features with those of VGG-Face by our proposed CM-R technique. The two models of ResNetShort and VGG-Face are trained from different face images by optimizing different loss functions through different architectures. This makes the learned discriminative information contained in ResNetShort features and VGG-Face features mutually complementary to each other thus better performance is achieved. The FR performance of CNNs drops fast when CNNs are applied to face images of low-resolution. Existing CNN methods can not deal with multiple resolutions of probe images. We propose a Deep Coupled ResNet model which extracts robust feature representations and multiple coupled mappings from face images to tackle the resolution degradation of probe images.

Publications

(i) **Ze Lu**, Xudong Jiang, and Alex Kot, "Feature fusion with covariance matrix regularization in face recognition." *Signal Processing*, 2018.

(ii) **Ze Lu**, Xudong Jiang, and Alex Kot, "A color channel fusion approach for face recognition." *IEEE Signal Processing Letters*, 2015.

(iii) **Ze Lu**, Xudong Jiang, and Alex Kot, "Deep coupled ResNet for low-resolution face recognition." Submitted to *IEEE Signal Processing Letters*, 2017.

(iv) **Ze Lu**, Xudong Jiang, and Alex Kot, "Color space construction by optimizing luminance and chrominance components for face recognition." Submitted to *Pattern Recognition*, 2017.

(v) **Ze Lu**, Xudong Jiang, and Alex Kot, "A novel LBP-based color descriptor for face recognition." *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017.

(vi) **Ze Lu**, Xudong Jiang, and Alex Kot, "An effective color space for face recognition." *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016.

(vii) **Ze Lu**, Xudong Jiang, and Alex Kot, "Enhance deep learning performance in face recognition." *IEEE International Conference on Image, Vision and Computing (ICIVC)*, 2017.