Hybrid feature extraction of multimodal images for face recognition

Usman Cheema* and Seungbin Moon*
*Dept. of Computer Engineering, Sejong University

Abstract

Recently technological advancements have allowed visible, infrared and thermal imaging systems to be readily available for security and access control. Increasing applications of facial recognition for security and access control leads to emerging spoofing methodologies. To overcome these challenges of occlusion, replay attack and disguise, researches have proposed using multiple imaging modalities. Using infrared and thermal modalities alongside visible imaging helps to overcome the shortcomings of visible imaging. In this paper we review and propose hybrid feature extraction methods to combine data from multiple imaging systems simultaneously.

1. Introduction

Facial recognition systems encounter various challenges in an open world environment including but not limited to illumination variation, pose variation, expression, aging, occlusion, disguise and noise [1]. Various algorithms are being developed to overcome these challenges and facial recognition systems have evolved to be robust for security and surveillance applications [2]. With the advancement of technology various imaging modality devices are now readily available at affordable cost. CCTV cameras predominantly employ infrared modality for adverse lighting conditions. Thermal cameras are capable of capturing heat signature thus can differentiate between human body and inanimate objects. More information about a scene can be captured using two or more imaging modalities. Current research is being done to combine image features or decision scores from different modalities for face recognition.

In this paper, we present a feature level fusion algorithm of face images taken simultaneously in visible and thermal spectra. Section 2 reviews the current methods in modality fusion. In Section 3 we present our feature fusion methodology followed by concluding remarks and future work in Section 4.

2. Multi-modal fusion

Various degrees of success has been achieved in face recognition using visible, thermal, and infrared images. Improved performance has been reported using multi-modal approaches for identity recognition. Multi-modal approaches use two or more modality images, e.g. infrared + visible, thermal + visible, as input. Multi-modal approaches help over come the short comings of each modality such as thermal images are sensitive to ambient temperature, visible images are sensitive to illumination etc.

Fusion of two or modalities is performed at feature level, or score level.

Feature level fusion

In feature level fusion, separate features are extracted from all images and combined to create hybrid features. These hybrid features are further used by a classification algorithm for facial recognition.

Score level fusion

In score fusion independent matching scores for the two modalities are calculated and the results are combined to obtain and final mean score.

3. Experimental Works

Dataset

Facial images from a proprietary disguise database, collected by the authors, were used as training and testing data. The data base is unique in the regard that it contains identical images of subjects captured using two different cameras, a visible imaging camera and a thermal camera.

The database also contains images of subjects wearing various disguise add-ons such as glasses, scarf, wig, makeup and fake and real facial hair. A sample set of images used are given in Figure 1.

Preprocessing

Three preprocessing schemes were applied: CSDN [3], SQI [4] and GIST [5] to refine the raw images for later stages.

Histogram equalization and difference of gradients was performed as second stage of preprocessing on all images. These methods have been proved to be effective for image normalization and local variation reduction in facial images [4,6].

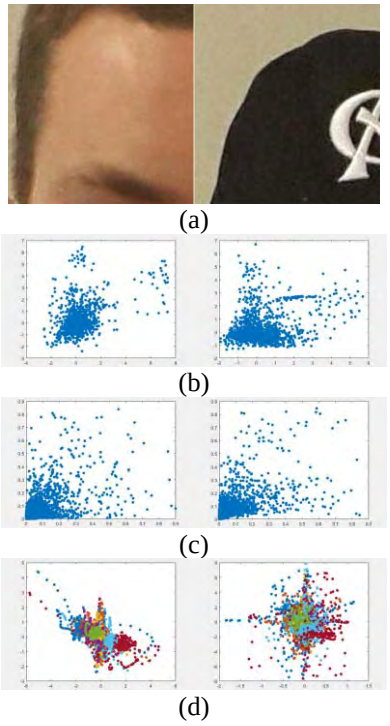
Hybrid Feature Extraction

We compare and present three feature level fusion techniques, Local Binary Patterns [7], Gabor wavelets [8] and Shearlets [9].

Image features for a visible image and its corresponding aligned thermal image are calculated. As each of these feature extraction methods extract features of a small image window, feature values for visible and thermal image are plotted against each other in a 2D plane for visualization purposes. Figure 2. shows a sample set of plotted hybrid features for sample images. The resultant hybrid features can be used in two ways; raw data or hybrid feature image as shown in Figure 2. The raw data forms a $n \times 2$ matrix where n is dependent on the image and feature extraction window size. The hybrid feature image can be used as a raw image input or its dimensions further reduced through a feature extraction method. The hybrid feature image can also be treated as a point cloud and its statistical distribution can be

used as features.

(Figure 1) Sample visible and thermal images



(Figure 2) (a) Original images, (b) Gabor hybrid features, (c) LBP hybrid features, (c) Shearlet hybrid features.

4. Conclusion and Future Work

Robust facial recognition systems are of increasing importance for security and access control. With the rising challenges of spoofing, facial recognition systems need to adapt stay ahead. Multi-modal facial recognition is promising approach against disguise and replay attack scenarios. In this paper we present hybrid feature calculation models, using various feature extraction techniques. The statistical analysis shows that the proposed hybrid features are promising and merit further work.

Various other feature extraction techniques such as HOG, PCA, SIFT can be combined using the same methodology and improved accuracies achieved. We also plan to perform disguise detection and facial recognition on visible and thermal facial images using the proposed hybrid feature calculation schemes.

References

[1] M. Ghayoumi, “A review of multimodal biometric systems: Fusion methods and their applications”, IEEE/ACIS 14th International Conference on Computer and Information Science (ICIS), 2015, p. 131-136.

[2] P. J. Phillips et al., “Overview of the face recognition grand challenge”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1 2005, p 947-954.

[3] B. Klare, and A.K. Jain, “Heterogeneous face recognition using kernel prototype similarities”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, 2013, p. 1410–1422.

[4] J. Choi, S. Hu, S.S. Young, and L.S. Davis, “Thermal to visible face recognition”, Proceedings of SPIE, 2012, p. 8371.

[5] A. Oliva, and A. Torralba, “Modeling the shape of the scene: A holistic representation of the spatial envelope”, International Journal of Computer Vision, vol. 42, 2001, p. 145–175.

[6] C. Chen, and A. Ross, “Matching thermal to visible face images using hidden factor analysis in a cascaded subspace learning framework”, Pattern Recognition Letters, vol. 72, 2016, p. 25-32.

[7] DC. He, and L. Wang, “Texture Unit, Texture Spectrum, And Texture Analysis”, IEEE Transactions on Geoscience and Remote Sensing, vol. 28, 1990, p. 509 - 512.

[8] Lee Tai S. “Image Representation Using 2D Gabor wavelets”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, 1996, p. 959-971.

[9] G. Kutyniok, W.-Q Lim, and R. Reisenhofer, “ShearLab 3D: Faithful Digital Shearlet Transform based on Compactly Supported Shearlets”, ACM Transactions on Mathematical Software, vol. 42, 2016 p. 5-42.