Deep Perceptual Mapping for Thermal to Visible Face Recognition

Teoh Soon Gi 1121116020
Kuan Wei Sern 1121116638

September 18, 2015

1 Executive Summary of Research Proposal

The objective of this research is using deep perceptual mapping (DPM) method improvement matching face in the night-time surveillance, so using the method to detect thermal to store a visible image of high resolution but the problem is the thermal- visible face recognition is due inheritance large modality difference and nature problem, to make systems in low correct match rate. The paper are trying to find what can make two modality closer, detector can recognition in real-time without computational overhead and what the effect of natural problem such as lighting/illumination, expression and identity. The mid-wave infra-red 'MWIR'-visible has a very less of pronounced modality gap and because of long-wave infra-red 'LWIR'-visible has a higher resolution so present DPM methods are harder. This article proposes a technique deep perceptual mapping to bridge the modality gap by a significant margin. With development of deep perceptual mapping for thermal to visible face recognition, the DPM is easy to train, very useful and able real-time.

2 Introduction

In thermal to visible face recognition facing a very common problems that is in the difference environment condition will deduced the performance of the face recognition. Its cant identify the face accuracy. The environment must only have spectrum will improve the performance, in converse, in the dark environment the system couldnt detect the face or have a low performance. So currently intra-red spectrum is the most stable to identify effects because its doesnt easier affected by any environments. The objects of the research in cross-modal face recognition to identify a person captured by intra-red spectrum and through the active infra-red sensors can perform in the night-time when the image is captured covertly. In intra-red spectrum, it can be without any light sources and it can be through the skin tissue to get thermal signatures and it is through passive thermal sensor. In intra-red spectrum can be divided by four main ranger that is Near (NIR), short-ware (SWIR), mid-ware (MWIR) and long-ware (LWIR) intra-red. NIR and SWIR are close the visible spectrum because they are reflection dominant but MWIR and LWIR are emission dominant so that they close to the thermal spectrum. Near intra-red and short-ware intra-red have a good results at the visible face matching because they have a small spectral gap and they modalities are quite similar.The limitation that is in the night-time surveillance,so it's only focus on thermal to visible.

3 Justification of Research

Thermal- to-visible facing a challenge matching problem because of there have a very large of modality gap. The research present a way to using significant margin to connect the modality gap. Using deep perceptual mapping capture the non-linear relationship and two modalities. The modality difference more

that 40% so in the thermal-visible face recognition using deep neural network to connect with the modality gap. Because of modalities and resolution in deep perceptual mapping are difference so the relationship between of them are non-linear so it is have the large gap between visible image and thermal.

4 Research Objectives

- To study technique of face recognition system for detection dark environment.

- To design face recognition technique to provide efficient and real time capable.

- To enhance face recognition system by using deep perceptual mapping for thermal to visible.

- To test the performance of face recognition system in University of Notre Dame (UND) collection dataset.

5 Literature Review

The first of compare studies is face matching between visible light images and near infrared (Yi, Liu, Chu, Lei, & Li, 2007). It is the early to explore the NIR to visible face recognition. They adopt canomical correspondence analysis and Latent Dirichlet allocation (LDA) to implement linear regression between the visible images and Near infra-red (NIR).

Visible to thermal face recognition (Socolinsky & Selinger, 2002) is one earlier studies carried out by Socolinsky and Selinger. They found out that "LWIR thermal imagery of human face not only a valid biometric, but almost surely a superior one to comparable visible imagery".

Some approach also focused on short-wave infra-red (SWIR)- to visible face recognition (Nicolo & Schmid, 2012) (Bourlai, Kalka, Ross, Cukic, & Hornak, 2010). SWIR or NIR to visible face matching carried out best result and similar modalities because of small spectral gap. The two approach only can adopt in night time surveillance application, for further research is needed in the thermal to visible matching domain.

6 Research Methodology

The deep perceptual mapping (Sarfraz & Stiefelhagen, 2015) can be divided into two parts:

I Deep Perceptual Mapping (DPM) Model

The deep perceptual mapping is based on a deep feed forward deep neural network (Glorot & Bengio, 2010). The target of training the deep network is to gain projections that can be used to bring two modalities closer. The simple feed forward neural network is giving a nice fitting architecture for such an objective. First provided the output of level hidden layer as:

$$H(x) = h^{(N)} = g(W^{(N)}h^{(N-1)} + b^{(N)}) \tag{1}$$

Assume that N+1 layers with m(k)in k-th layer, so the layer k=1,2,N, each layer has a non-linear output by using matrix W and non-linear function $g(\cdot)$. So input output of k-th become $h^{(k)} = g(W^{(k)}h^{(k-1)} + b^{(k)})$ and $b^{(k)}$ is a bias vector. The mapping H is a perceptual mapping function so with the output $H(x) = h^{(N)} = g(W^{(N)}h^{(N-1)} + b^{(N)})$ so equation 2 final output:

$$\overline{X} = s(W^{(N+1)}H(x)) \tag{2}$$

and 's' is linear mapping. We use the formula of equation 2 determine the parameters 'b' and 'W' because of objective function must minimize to differentiate between the thermal and visible. So formulate of DPM as:

$$arg\frac{min}{W,b}J = \frac{1}{M}\sum_{i=1}^{M}(\overline{X}_i - t_i)^2 + \frac{\lambda}{N}\sum_{k=1}^{N}(||W^{(k)})||_F^2 + ||b^{(k)}||_2^2) \tag{3}$$

The first term objective is to simple squared loss with the network and the corresponding training and the second terms is regularization term with the regularization parameter.

For DPM training, we adopted densely computed feature representation overlapping small region in the image. This can make model easy to taking the differing local region's perceptual, ease the need of large training image and show the input in approximately small dimension. After that, use same identity corresponding image assure the model can learn only show differences due to the modality as the other appearance parameters such as expression, identity, lighting and so on. The stochastic gradient descent (SGD) (Bottou, 2010) approach is used for updated the standard back projection of error from computing the gradient of the loss in each iteration. We adopt hyperlic tangent 'tanh' function for non linear activation function g (z) in invisible layers.

II Thermal - Visible Face Matching

We used same approach as (Choi, Hu, Young, & Davis, 2012) to matching face. we use ideal for the surveillance scenario, so the gallery image can store offline and processed during test period no transformation and overhead is necessary. Furthermore, the detector able to recognition in real-time using without any computational overhead. The presented DPM is not dependent on the direction of mapping. We examined the opposite thermal to visible only slight performance variation.

The authors have done few test to evaluate the efficient of the DPM system. The first test result show the system can make accuracy rate of 83% when a comparing thermal image a varied number of well-lit images. 60% when compare two available/subject. And 55% when doing one to one comparison. All the test is using University of Notre Dame's (UND) collection X1.

References

Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proceedings of compstat'2010* (pp. 177–186). Springer.

Bourlai, T., Kalka, N., Ross, A., Cukic, B., & Hornak, L. (2010). Cross-spectral face verification in the short wave infrared (swir) band. In *Proceedings of the 2010 20th international conference on pattern recognition* (pp. 1343–1347). Washington, DC, USA: IEEE Computer Society. Retrieved from `http://dx.doi.org/10.1109/ICPR.2010.1115` doi: 10.1109/ICPR.2010.1115

Choi, J., Hu, S., Young, S. S., & Davis, L. S. (2012). *Thermal to visible face recognition* (Vol. 8371). Retrieved from `http://dx.doi.org/10.1117/12.920330` doi: 10.1117/12.920330

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *In proceedings of the international conference on artificial intelligence and statistics (aistats10). society for artificial intelligence and statistics.*

Nicolo, F., & Schmid, N. (2012, Dec). Long range cross-spectral face recognition: Matching swir against visible light images. *Information Forensics and Security, IEEE Transactions on, 7*(6), 1717-1726. doi: 10.1109/TIFS.2012.2213813

Sarfraz, M. S., & Stiefelhagen, R. (2015). Deep perceptual mapping for thermal to visible face recognition. *CoRR, abs/1507.02879*. Retrieved from `http://arxiv.org/abs/1507.02879`

Socolinsky, D. A., & Selinger, A. (2002). A comparative analysis of face recognition performance with visible and thermal infrared imagery. In *Proceedings of the 16 th international conference on pattern*

recognition (icpr'02) volume 4 - volume 4 (pp. 40217–). Washington, DC, USA: IEEE Computer Society. Retrieved from http://dl.acm.org/citation.cfm?id=846227.848532

Yi, D., Liu, R., Chu, R., Lei, Z., & Li, S. Z. (2007). Face matching between near infrared and visible light images. In *Proceedings of the 2007 international conference on advances in biometrics* (pp. 523–530). Berlin, Heidelberg: Springer-Verlag. Retrieved from http://dl.acm.org/citation.cfm?id=2391659.2391719

Tittle of research project	Deep Perceptual Mapping for Thermal to Visible Face Recognition
Member of project	1. Teoh Soon Gi 1121116020 2. Kuan Wei Sern 1121116638
Executive Summary (5 marks)	
Introduction (3 marks)	
Justification of research (3 marks)	
Research objective (3 marks)	
Literature review (6 marks)	
Research Methodology (8 marks)	
References (2 marks)	