Australian Journal of Basic and Applied Sciences, 9(23) July 2015, Pages: 522-527

ISSN:1991-8178

Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com

Template Matching in Supervised Learning for Thermal Image Classification

W.S. Lim, Isabella N. and B.C. Yeo

Faculty of Engineering and Technology, Multimedia University, Jalan Ayer Keroh Lama, Melaka, Post code 75450, Malaysia

ARTICLE INFO

Article history:
Received 12 March 2015
Accepted 28 June 2015
Available online 22 July 2015

Keywords:
Sum of Absolute Differences, Structural Similarity Index Metric, Probabilistic Neural Network, template matching, thermal imaging, image classification.

ABSTRACT

In this paper, the precision and feasibility of Structural Similarity Index Metric (SSIM) in labelling training data for a supervised neural network has been investigated. This proposed model of data-labelling with an expert system has been applied in thermal image classification. The well-known Sum of Absolute Differences (SAD) has also been adopted for performance comparison in template matching process, and the Probabilistic Neural Network (PNN) is chosen as the image classifier. The experimental result shows that SSIM produces higher matching accuracy than the SAD technique, and subsequently provides a more precise set of training data to supervise the neural network in classifying the thermal images. By evaluating the PNN performance using data labelled by both SAD and SSIM techniques, it is further proven that SSIM facilitates the classifier in achieving higher classification accuracy. The outcome of this research has shown that SSIM serves as a superior template matching method in generating labelled data for training a supervised neural network like PNN.

To Cite This Article: W.S. Lim, Isabella N. and B.C. Yeo., Template Matching in Supervised Learning for Thermal Image Classification. ***Aust. J. Basic & Appl. Sci.,*** *9(23): 522-527, 2015*

INTRODUCTION

Image classification which lies in the "high end" of the image analysis field is a common problem and has long since been researched. It can be defined as the task of assigning an input image one label from a fixed set of categories and involves image processing as a pre-classification step. It incorporates diverse algorithms and has a wide range of applications which include; traffic pattern analysis, medical image diagnosis, bar code reading and face or finger print classification (Gao, Y., 2012).

Template matching entails evaluation of the similarities between a test image and a set of templates/references using similarity measures. A test image with the highest similarity to any of the templates makes it of the class represented by the template. This approach is conceptually simple. The two similarity measures employed in this specific study are Sum of Absolute Differences (SAD) and Structural Similarity Index (SSIM). Traditionally, SAD measures the prominence of errors betwixt a test image and the representative image using a range of recognised characteristics of the human visual system (HVS). Based on the hypothesis that human visual apprehension is greatly conformed to selecting structural data from a scenery, SSIM Index, a more recent alternative image similarity measure was initiated based on the decomposition of structural data. It is said that the extent of structural detail alternation can render a good estimation of apprehended image degradation (Mai, Z.Y., 2006; Wang, Z., 2004). The effectiveness of the template matching method relies on the image similarity measure being employed.

Arthur Samuel outlined in 1959 that: "[Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed." Machine learning simulations identify profiles in data and change program decisions consequently (McCrea, N., 2015). Two kinds of machine learning are in existence; supervised learning and unsupervised learning. In supervised learning under which our interest PNN lies, an inflow vector is portrayed via the inputs together with a set of targeted outcomes, one for each artificial neuron. Pertaining to the output phase, a forward pass is carried out, and the disparities between the targeted and true outcome per node observed. These are thereafter employed in choosing weight adjustments in the net following the predominant learning rule. The term supervised is drawn from the concept that the targeted signals on particular output nodes are contributed by an outside teacher.

It is therefore essential that the data used to train the PNN rightly represents each class. The aim of

Corresponding Author: W.S. Lim, Faculty of Engineering and Technology, Multimedia University, Jalan Ayer Keroh Lama, Melaka, Post code 75450, Malaysia
E-mail: wslim@mmu.edu.my

this paper is to study the image labelling capability of; two similarity measures SAD and SSIM. These labeled images from each similarity measure are thereafter being used for training the PNN and its feasibility is tested. The images under test for this study were obtained using a thermal camera.

Thermal imaging (thermography) can be defined as a proficiency for generating a depiction of infrared light released by objects with the aid of a thermal camera. The thermal imaging cameras work by capturing the infrared (IR) energy transmission from a physical item to its surrounding and creating a time-use image where warmer objects are reckoned brighter and cooler objects are reckoned darker by default. All objects above absolute zero are emitters of radiation. Each of the thermograms has its own temperature and therefore can be distinguished easily. Thermal imaging was opted for instead of other imaging techniques for example CCTV cameras in this case because of its advantages which include: full awareness, being weatherproof, low maintenance, extreme difficulty to escape detection since thermal contrast is impossible to conceal and not being affected by day and night cycle, smoke, fog and haze (Cui, H., 2013).

Image Labelling and Classification Methods:

There are two similarity measures; Sum of Absolute Differences (SAD) and Structural Similarity Index Metric (SSIM) being discussed below in the template matching method earlier introduced. They work on the basis that the similarities between a test image and a set of templates/references are obtained after which the class label is determined based on the best match of each test image with any of the templates in terms of similarity. Both are full reference methods meaning that a complete reference image is assumed to be known.

A. Sum of Absolute Differences (SAD):

SAD by default is a dissimilarity measure that adds up the absolute differences between elements/pixels in the template macro block and the test macro block. The variations are added up to generate a single measure of similarity which is equivalent to the L1-norm of the difference, also known as Manhattan- or Taxicab-norm (Han, J., 2011). For example, let us assume that a 2-D MxN reference, A(x, y) is to be paired within a source frame B(x, y) of dimensions PxQ whereby (P>M and Q>N). For every pixel position (x, y) in the frame, the SAD value is computed as below (Alsaade, F., 2012):

$$SAD(x,y) = \sum_{k=0}^{(M-1)(N-1)} \sum_{l=0}^{(N-1)} \left| S(x+k, y+l) - T(k,l) \right| \tag{1}$$

In this algorithm, the smaller the distance measure obtained by SAD function between the reference image A and a sub-image in the source frame B, the closer match between the searched template and that corresponding sub-image is. The range of the SAD value varies from 0 to 1.Therefore, if the measured distance by SAD function is zero/close to zero, the local sub-image is similar/somewhat similar to the reference and if the measured distance by the SAD function is one/close to one, the local sub-image is dissimilar /somewhat dissimilar to the template.

The SAD algorithm is extensively used, simple, has good accuracy, and has little complexity in implementation to find a relationship between the image windows. Unfortunately, it has been said to produce misleading values that do not correlate perfectly with the human perceived measurement from past studies (Zheng, Y., 2012).

Given the mentioned drawbacks of the Sum of Absolute Differences (SAD), a more intelligent solution to the image quality assessment and comparison namely; the SSIM algorithm was proposed. Its basis is on separated relations of local luminance, contrast and structure between a reference image and a test image. Previous experiments show that the SSIM index technique, which can be employed simply, has a higher possibility of matching with Human Visual System (HVS) than SAD which this project sets out to investigate.

SSIM index establishes the value of structural information decomposition which is composed of three relations; luminance, contrast and structure between the reference image and an image. These comparisons can be better be defined as below as presented by Gao (2012).

Given two local image patches from the original and comparison image respectively:

B. Structural Similarity Index Metric (SSIM):

$$x = \{x_i | i = 1, 2, \ldots, M\} \quad and \quad y = \{y_i | i = 1, 2, \ldots, M\} \tag{2}$$

The local SSIM is:

$$(x, y) = l(x, y), c(x, y), s(x, y) \tag{3}$$

Where $l(x, y)$ is Luminance comparison, $c(x, y)$ is Contrast comparison and $s(x, y)$ is Structure comparison. These three terms are defined as:

$$l(x,y) = \frac{2\mu_x \mu y + C_1}{\mu_x^2 + \mu y^2 + C_1} \quad (4)$$

$$c(x,y) = \frac{2\sigma_x \sigma y + C_2}{\sigma_x^2 + \sigma y^2 + C_2} \quad (5)$$

$$s(x,y) = \frac{\sigma xy + C_3}{\sigma_x \sigma y + C_3} \quad (6)$$

and y singly represent the mean intensity of x and y in the order given, and y singly stand for the standard deviation of x and y in that order given and symbolises the covariance of x and y. $C_1 = (K1L)^2$, $C_2 = (K2L)^2$, $C_3 = C_2 / 2$ are constants that have been identified as of significance in describing the saturation impacts of the optical system where there is small luminance and contrast areas and balance the working of the metric when the denominators approach zero.

The local SSIM is thus:

$$SSIMl(local) = \frac{(2\mu_x \mu y + C_1)(2\sigma_x \sigma y + C_2)}{(\mu_x^2 + \mu y^2 + C_1)(\sigma_x^2 + \sigma y^2 + C_2)} \quad (7)$$

When this local evaluation is employed in the analysis an entire image using a sliding window technique, an SSIM quality map is produced. The final SSIM metric of the entire image is basically the mean of the SSIM map. In addition, the higher the value of SSIM (x, y) is, the more similar the image x and y are.

C. Supervised Learning Method:

Initially suggested by Donald Specht in 1990, the Probabilistic Neural Network (PNN) model follows the working of supervised learning. It is a feed forward artificial neural net which means data is fed into the inflows and moves through the network, stage by stage, till it reaches the outflows with no feedback between the stages It is also modelled on a statistical method called Kernel Fisher discriminant analysis (Mika, S., 1999) and is employed in nonlinear evaluation which approximates the Bayes optimal decision limits by predicting the probability density function of the targeting dataset by employing the Parzen nonparametric estimator. The PNN structure consists of four stages or layers: input stage, pattern stage, summation stage, and decision stage.

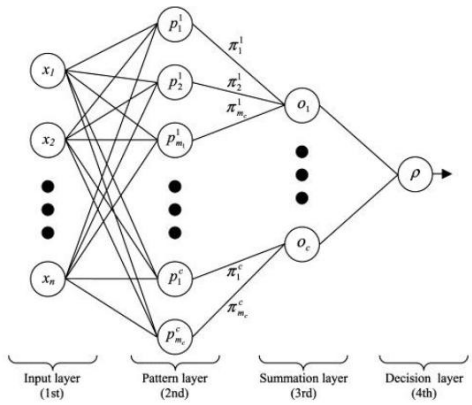

Fig. 1: Probabilistic Neural Network Architecture.

The above figure shows a PNN structure. The first stage portrays the input characteristics. The magnitude of artificial neurons in the pattern stage is equivalent to the magnitude of training occurrences. The magnitude of artificial neurons in the summation stage is equivalent to the magnitude of categories in the training occurrences. The input stage is completely linked to the pattern stage. The input stage does not carry out any calculation and plainly allocates the input to the nodes in the pattern stage. The pattern stage is semi- linked to the summation stage. Each set of training occurrences matched to each class is just connected to one artificial neuron in the summation stage. That is, the summation components basically add up the inputs from the pattern components that are in accordance with the class from which the training data was chosen. Further information on the working of the PNN is also presented in (Ganesula, K., 2010; Mao, K.Z., 2002).

Experimental Setup:

A sequence of video frames were obtained by road video surveillance taken using a thermal camera. Image processing was also carried out particularly; background subtraction with the implementation of the foreground detector based on Gaussian mixture models (GMMs) to extract the moving objects, edge detection and cropping. MATLAB programming simulation tool was implemented in this study.

The process of labelling and classifying the processed images into the three categorical classes; adjoining cars, trees and cars entailed finding the reference images per class from 10 images belonging to each class and then using these reference images

to label a set of forty test images belonging to each of the three categorical classes. This was done based on each of the two similarity measures SAD (Sum of Absolute Differences) and SSIM (Structural Similarity) Index and their labelling accuracy was noted. A half of the forty test images used were created by adding noise to the twenty original test images. Consequently, the PNN was trained using the set of labelled images and tested using a fresh sample of images different from the ones used in training. This training and testing process was done for each of the two similarity metrics; SAD and SSIM Index labelled images. Their classification accuracy was tested with the same fresh sample of images and noted.

A. Finding the reference image:

The reference images representing the three classes were obtained using each of the two similarity measures SAD and SSIM following the procedure below:

Sum of Absolute Differences (SAD) measure:

In order to find the averaged/template image these are steps that were to be followed; of the mentioned 10 images belonging to each class, for each class, choose an image as a reference image and obtain its SAD with the rest of the images, sum all the SADs of that outcome and the average SAD value corresponding to that reference image can be calculated and recorded. All the images in the class take turn to be the reference image and the process is repeated. The reference image that produces the lowest average SAD value is chosen as the class template. Following this, all three class templates pertaining to SAD can be found.

Structural Similarity Index Metric (SSIM) measure:

In order to find the averaged/template image these are steps that were to be followed; of the mentioned 10 images belonging to each class, for each class, choose an image as a reference image and obtain its SSIM Index with the rest of the images, sum all the SSIM Index values of that outcome and the average SSIM Index value corresponding to that reference image can be calculated and recorded. All the images in the class take turn to be the reference image and the process is repeated. The reference image that produces the highest average SSIM value is chosen as the class template. Following this, all three class templates pertaining to SSIM can be found.

Resulting template images:

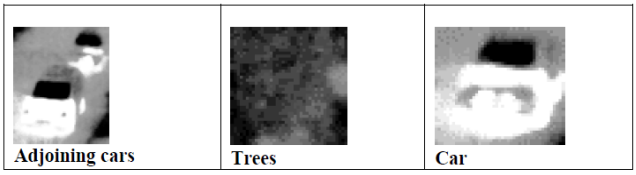

Fig. 2: SAD template images.

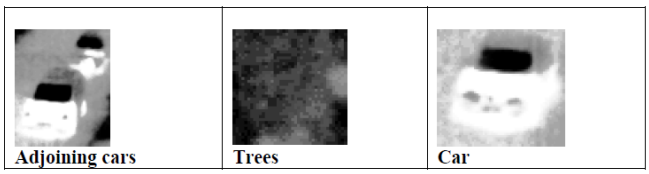

Fig. 3: SSIM template images.

B. Labelling the images:

The images were labelled using the two similarity measures SAD and SSIM in template matching following the procedure below:

Sum of Absolute Differences (SAD) measure:

In order to label a test image, the Sum of Absolute Difference between each of the forty test images in each class and all three class template images obtained by the SAD technique is calculated. The class reference that results into the lowest SAD with any test image is noted and it is concluded that the test image belongs to that class as represented by the reference image. Under normal circumstances, a test image belonging to any class should produce the lowest SAD with the reference image representing that same particular class. If this does not hold, it is an indicator that the test image has been wrongly classified.

Structural Similarity Index Metric (SSIM) measure:

In order to label a test image, Structural Similarity Index between each of the forty test image in each class and all three class template images obtained by the SSIM technique is calculated. The class reference that results into the highest SSIM with any test image is noted and it is concluded that the test image belongs to that class as represented by the reference image. Under normal circumstances, a test image belonging to any class should produce the highest SSIM index with the reference image representing that same particular class. If this does

not hold, it is an indicator that the test image has been wrongly classified.

C. Classification using Probabilistic Neural Network (PNN):

The PNN is then trained in turn using labelled images from each of the similarity measures and testing is done for each of the similarity metric labelled PNNs using a fresh sample of images different from the ones used in training. If a test image returns a prediction different from the one representing the class it belongs to, it can be concluded that the test image has been wrongly classified. The corresponding accuracy is noted.

RESULTS AND DISCUSSION

A sequence of video frames were obtained by road video surveillance taken using a thermal camera and processed. Table I depicts the accuracy of the two template matching methods and Table II shows the classification accuracy of PNN.

Table I: Labelling Accuracy of Template Matching Methods.

	Sum Of Absolute Differences (SAD)	Structural Similarity Index Metric (SSIM)
Adjoining Cars	70%	82.5%
Trees	95%	95%
Cars	92.5%	94%
Average	85.8%	90.5%

Table II: Classification Accuracy Of Pnn Trained By Different Template Matching Methods.

	PNN trained by SAD	PNN trained by SSIM
Adjoining Cars	70%	90%
Trees	100%	100%
Cars	50%	98%
Average	73.3%	96.0%

***Performance study in labelling by SAD and SSIM*:**

Both the similarity measures SSIM and SAD label the trees class images with high accuracy at 95%. This shows that both methods are competent as the trees class can be well distinguished from the adjoining cars class and the cars class. The SAD labelled the cars class images with 92.5% accuracy whereas 94% for SSIM, placing a few images into the adjoining cars class. However, the adjoining cars class images are labelled with 70% accuracy for the SAD measure and 82.5% accuracy for the SSIM measure. This shows that it is tricky for both measures to rightly label the adjoining cars class images and most of them are wrongly placed in cars class. Nonetheless, the SSIM measure stands out labelling the adjoining cars class with higher accuracy as compared to the SAD. The overall precision of SAD in labelling the images for the three classes is 85.8% while that of SSIM is 90.5%.

***Performance study in classification with the PNN*:**

The PNN produces 100% accuracy in classification of the trees class images when testing is carried out after it has been trained with both SSIM and SAD labelled images. However, for the classification of the adjoining cars, subsequent to training the PNN with SAD labelled images, testing produces accuracy of 70% while for the SSIM labelled images, the accuracy is 90%. Also, for the classification of the cars, subsequent to training the PNN with SAD labelled images, testing produces accuracy of only 50% while for the SSIM labelled images, the accuracy is 98%. On average, the PNN achieved much higher accuracy at 96% when trained with images labelled by SSIM compared to 73.3% in SAD labelling method.

This shows that SSIM is more efficient than SAD in labelling the training images employed in the PNN and in the long run ensures higher accuracy of the PNN for classification in contrast to SAD.

This could be attributed to the fact that while the SAD estimates perceived errors; on the other hand, SSIM takes into account image degradation as perceived change in structural information. Structural information is the idea that the pixels have strong inter-dependencies especially when they are spatially close. These dependencies carry important information about the structure of the objects in the visual scene.

Conclusion:

The Sum of Absolute Differences shows its capability in image classification but can be tabbed as average since when employed in labelling the training images for the PNN, it leads to only moderate accuracy in differentiating the three classes. However, when compared to SSIM labelled images in training, higher accuracy is portrayed making Structural Similarity Index a much better similarity metric in labelling for the PNN. In other words, the SSIM acts superiorly in providing more precise set of training data to supervise the neural network in classifying the thermal images. It facilitates the classifier with a more informative

learning process that resulted in higher classification accuracy. The outcome of this research has shown that SSIM serves as a superior template matcher in generating labelled data for training the supervised neural network PNN.

REFERENCES

Alsaade, F., 2012. "Fast and Accurate Template Matching Algorithm Based on Image Pyramid and Sum of Absolute Difference Similarity Measure," Research Journal of Information Technology, vol. 4, no. Academic Journals Inc., pp: 204-211.

Cui, H., Y. Xu, J. Zeng, Z. Tang, 2013. “The Methods in Infrared Thermal Imaging Diagnosis Technology of Power Equipment,” Electronics Information and Emergency Communication (ICEIEC).

Ganesula, K., 2010. "Probabilistic Neural Networks," CPSC 636: Neural Networks Spring.

Gao, Y., 2012. "Complex-Wavelet Structural Similarity Based Image Classification," Master's Thesis, Waterloo, Ontario, Canada.

Han, J., M. Kamber and J. Pei, 2011. “Data Mining: Concepts and Techniques,” Data Mining Elsevier.

Mai, Z.Y., C.L. Yang, K.Z. Kuang and L.M. Po, 2006. "A Novel Motion Estimation Method Based On Structural Similarity For H.264 Inter Prediction," in IEEE Acoustics, Speech and Signal Processing, Toulouse.

Mao, K.Z., K.C. Tan and W. Ser, 2002. "Probabilistic Neural Network Structure Determination for Pattern Classification," IEEE Neural Networks.

McCrea, N., 2015. “An Introduction to Machine Learning Theory and Its Applications: A Visual Tutorial with Examples,” Toptal.

Mika, S., 1999. "Fisher Discriminant Analysis with Kernels," in Neural Networks for Signal Processing IX, 1999. Proceedings of the 1999 IEEE Signal Processing Society Workshop.

Wang, Z., A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, 2004. "Image Quality Assessment: From Error Visibility to Structural Similarity," in IEEE Trans. Image Processing, 13(4): 600-612.

Zheng, Y., W. Zhu, P. Chen, 2012. "H.264 Visual Perceptual Coding in Uniform Analysing and Encoding Framework," Image and Signal Processing (CISP), 5th International Congress.