A NEW DISTANCE MEASURE EMPLOYING ELEMENT-SIGNIFICANCE FACTORS FOR ROBUST IMAGE CLASSIFICATION

Kunio Kawahara and Tadashi Shibata

Department of Frontier Informatics, School of Frontier Sciences, The University of Tokyo
701 Kiban Bldg. 5-1-5 Kashiwanoha, Kashiwa, Chiba, 277-8561, Japan
kawahara@else.k.u-tokyo.ac.jp, shibata@ee.t.u-tokyo.ac.jp

ABSTRACT

A new simple distance measure has been proposed in which each vector element is weighted in the distance calculation according to its importance as determined by taking its statistics into account. In order to reflect the characteristics of the class, the element-significance factors are calculated based on intraclass variances and mean values of vector elements and utilized in the distance measure. The proposed distance measure has been applied to the face detection system and the cephalometric landmarks identification system which we developed in other work. Improved performances in image classification have been demonstrated.

1. INTRODUCTION

Vector quantization (VQ) is widely used in various applications such as data compression systems and image recognition systems. VQ is such a technique that an input vector is mapped into the class which has the maximum-likelihood code vector to the input vector. In the process of vector matching, some types of distance measures are utilized to evaluate the dissimilarity between the input vector and code vectors. Euclidean distance and Manhattan distance are widely used due to their simplicity.

In VQ-based image recognition systems, the feature vectors are generated by extracting the characteristic features from images. If two images are similar to each other, two feature vectors generated from them are expected to be mapped closely in the vector space. Therefore, the distance measure plays an important role in classifying images in the vector space. In the vector representation, it is often the case that the relative importance of each vector element varies. Namely, if some vector elements represent the critical features of an image, then the elements are more important than others. In such a case, the relative importance among elements must be taken into account in the distance measure. However, the conventional Euclidean distance and Manhattan distance do not work for this purpose because they treat all elements with an identical weight. In order to accommodate to the issue, weighted distance measures have been proposed, where both interclass and intraclass variances in vector elements are utilized to determine the element weight factors in distance calculation [1, 2].

The purpose of this paper is to propose a new weighted distance measure where only the intraclass statistical characteristics are taken into account to determine the weight factors in the distance calculation. Only the variances and mean values of vector elements within a certain class of sample vectors are utilized to determine the weight factors. Therefore, the procedure is very simple. We applied the proposed measure to the VQ-based face detection system [3, 4] and to the cephalometric landmarks identification system [5], and the improved performances have been demonstrated.

2. NEW DISTANCE MEASURE

In this section, how the weighting factors are determined from the statistical distribution of samples. In this paper, Manhattan distance :

$$\sum_{i=1}^{n} |x_i - t_i|,$$

is utilized as the basis of the new weighted distance measure. Here x_i and t_i are the i-th elements of the input vector and the template vector (code vector), respectively, and n is the vector dimension. In the following, we introduce two important parameters : the significance index s_i and the significance factor S_i. s_i is an index related to the significance of the i-th vector element and is derived from the statistical data of t_i, i.e. from its standard deviation σ_i and mean value μ_i. S_i is a factor that determines the weight in the distance measure and is obtained by performing a non-linear transformation on s_i.

2.1 Concept of Significance Index

Since the features of a certain class exist in its own statistics, we define s_i in terms of only the standard deviation σ_i and the mean values μ_i of the element t_i within the class.

Now, we wish to define the meaning of "class" more clearly. Consider the problem of classifying input images into the class of faces or non-faces as in Refs [3, 4]. Within the class of faces, there exist a number of groups each representing a particular type of faces like round faces, long faces, etc. The statistical characteristics may vary from one group to another. Therefore, the significance index is determined for each group and not for the entire class of faces. Then, index "i" refers to each group in the class. In order to form groups within the same class, we employed the k-means clustering algorithm. In the following, σ_i and μ_i are calculated for each cluster (group) in the class. Hereafter, the term "class" refers to a cluster or a group as defined in this paragraph.

Next, we consider what determines the relative importance of elements in vectors belonging to the same cluster. The significant elements are those that represent the characteristic features of the cluster. Therefore their values would be stable. In other words, its variance would be small. We propose a basic assumption that σ (standard deviation of an element) is the primary index representing its significance. Namely, the smaller the σ is, the more significant the element is. However, σ alone is not a good significance index because σ tends to increase as the mean (μ) gets larger. Such

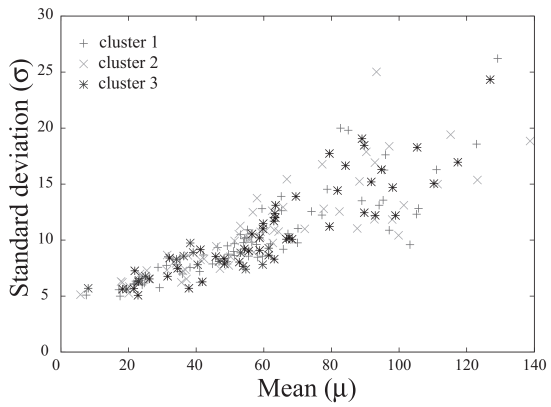


Fig. 1. Correlation between the standard deviation and the mean of vector element. These data are shown for three centroid vectors from clusters which were deduced from 300 face template vectors using K-means algorithm. One vector (PPED vector [5]) has 64 elements.

Table 1. Significance of vector element as related to the combination of standard deviation (σ) and mean (μ).

σ	μ	significance
small	large	more significant
small	small	average
large	large	average
large	small	less significant

a relationship is visible in Fig. 1. However, the point is that a large scattering of data is observed in the nearly-linear relation between σ and μ.

Our basic idea is that the very spread of data in Fig. 1 indicates the significance index of each element. If an element has a σ value smaller than the average of σ corresponding to the μ value, the element is regarded as significant because its statistical distribution is tighter than the average. On the contrary, if σ is larger than the average, then the element is regarded as less significant. Therefore the quantity :

$$s_i \equiv \frac{\sigma_i}{\mu_i},$$

can be a reasonable candidate for the significance index of i-th element. This represents the significance normalized to the mean value. This idea is illustrated in Table 1 which shows how the values of σ and μ are related to significance.

We also introduce another definition of significance index as :

$$s_i' \equiv \sigma_i(\frac{\sigma_i}{\mu_i}) = \frac{\sigma_i^2}{\mu_i}.$$

Here, σ_i is the primary index that represents the significance of the element. And the normalized significance σ_i/μ_i is multiplied on σ_i in order to take the geometrical average of these two indexes. This will strengthen its meaning to better represent the significance of the element. In Fig. 2, the two proposed significance indexes, s_i and s_i', are plotted as functions of σ. In (a), s_i decreases slightly with decreasing σ down to around 7–8, but increases below that point. This is in contradiction to our basic assumption that σ is the primary

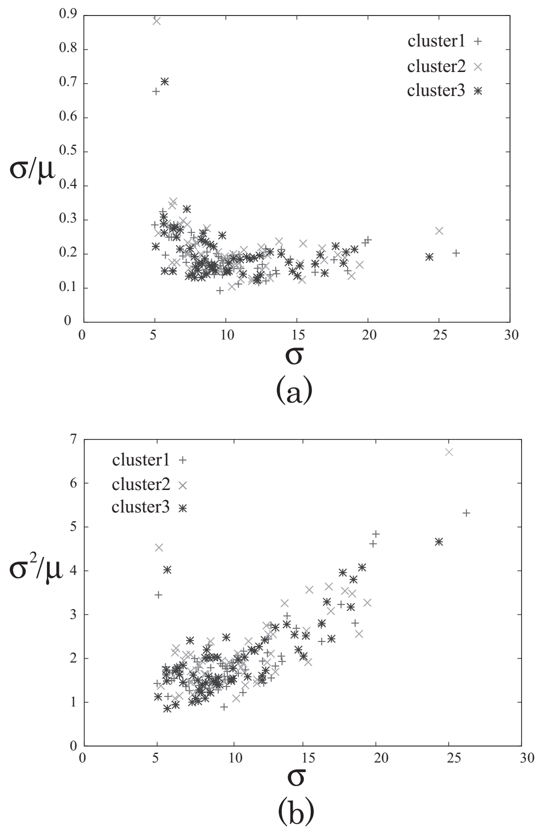


Fig. 2. Proposed two significance indexes s_i (a) and s_i' (b) as functions of σ.

significance index. Since larger values of σ and s_i mean they are less significant, s_i must monotonically decrease with σ. However, in (b) a good monotonic relation is seen between s_i and σ. These observations suggest that the definition of s_i' would be a better choice as the significant index. We will demonstrate this by experiments.

2.2 Significance Factor S_i Determined from Significance Index s_i

Now, we introduce standardized indexes for s_i (s_i') as in the following :

$$\hat{s}_i = -\frac{s_i - \bar{s}}{\Sigma_s}, \quad \hat{s}_i{}' = -\frac{s_i' - \bar{s}'}{\Sigma_{s'}}.$$

Here $\bar{s}$ ($\bar{s}'$) and Σ_s ($\Sigma_{s'}$) are the mean ant the standard deviation of s_i (s_i'), respectively. We put minus signs to indicate a larger $\hat{s}_i$ (s_i') means it is more significant.

In order to correlate the standardized significance index $\hat{s}_i$ ($\hat{s}_i{}'$) to weight factors of elements, we introduce a non-linear transformation to $\hat{s}_i$ ($\hat{s}_i{}'$). The sigmoidal function is employed as in Eq. (1) for this purpose :

$$S_i = \frac{1}{1+e^{-p\hat{s}_i}}, \quad S_i' = \frac{1}{1+e^{-p\hat{s}_i{}'}}, \tag{1}$$

where p is a parameter to control the slope of the function, as shown in Fig. 3. We call S_i and S_i' the significance factors which are used to determine the weight factors in the distance measure.

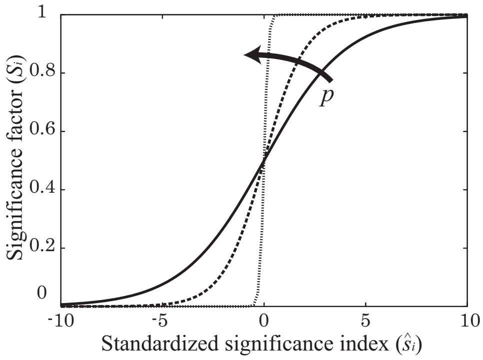


Fig. 3. Sigmoidal function used to transform s_i to S_i.

2.3 Proposed Weighted Distance Measure

The new weighted distance measure proposed in this work is defined using the significance factor S_i (S_i') as in Eq. (2) :

$$d \equiv \sum_{i=1}^{n} \frac{|x_i - t_i|}{\sigma_i} S_i \quad \text{or} \quad d \equiv \sum_{i=1}^{n} \frac{|x_i - t_i|}{\sigma_i} S_i'. \qquad (2)$$

$|x_i - t_i|$ is divided by σ_i because the contribution from each element difference to the total distance need be equalized. The difference between the elements must be evaluated taking its statistical characteristics into account. Therefore the difference in the element values normalized to the standard deviation σ_i, we believe, is a reasonable choice in this regard.

3. EXPERIMENTAL RESULTS AND DISCUSSIONS

The proposed new weighted measure was applied to two systems, the face detection system [3] and the cephalometric landmarks identification system [5]. As a vector generation algorithm, we employed Projected Principal-Edge Distribution (PPED) [5] which uses principal direction edges extracted from an image. In this algorithm, a 64x64-pixel gray-scale image is converted to a vector. The distances between the input vector and the centroid vectors of the clusters are calculated. In this calculation, only centroid vectors of clusters are utilized as template vectors since the proposed distance measure already contains the features of the cluster. Therefore, the number of the clusters is equal to the number of vectors actually used in the template matching. In this respect, the number of template vectors has been reduced drastically by using the new distance measure.

3.1 Face Detection System

In the face detection system developed in Ref. [3], face images and non-face images are utilized as template images. If the vector of an input image is closer to a template vector in the face images than those in the non-face images, the area of the image is regarded as a face area. The experimental results are demonstrated in Fig. 5. The results are compared for two cases, one with the weighted distance measure using s_i $(= \sigma_i/\mu_i)$ and one with the weighted distance measure using s_i' $(= \sigma_i^2/\mu)$. Here the number of clusters was chosen as 10. The average number of face samples in each cluster was about 30 which is sufficient to derive statistical parameters. In these results, only face segmentation is carried out and no verification processing is conducted. This is why so many false positives are still remaining.

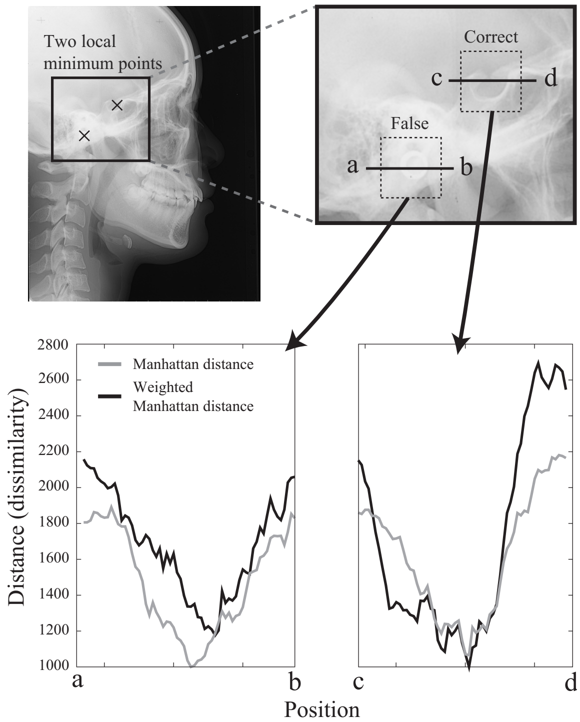
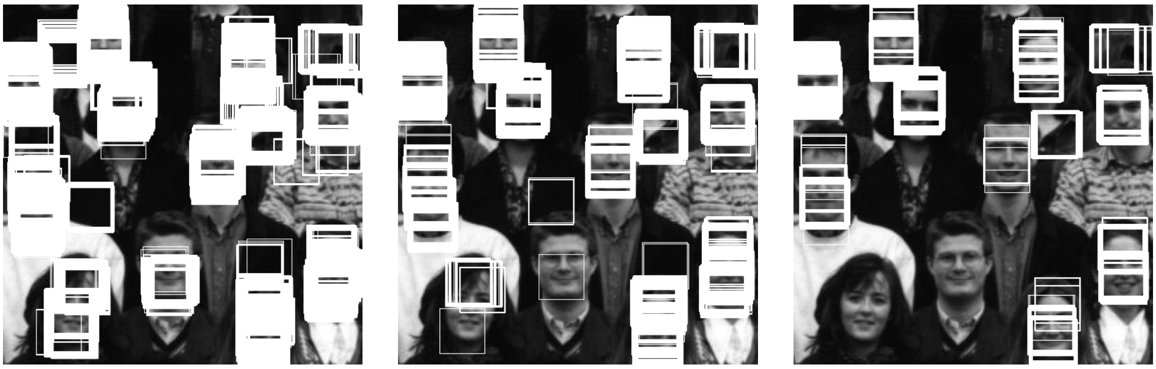


Fig. 4. Results of landmark identification with Manhattan distance and weighted distance.

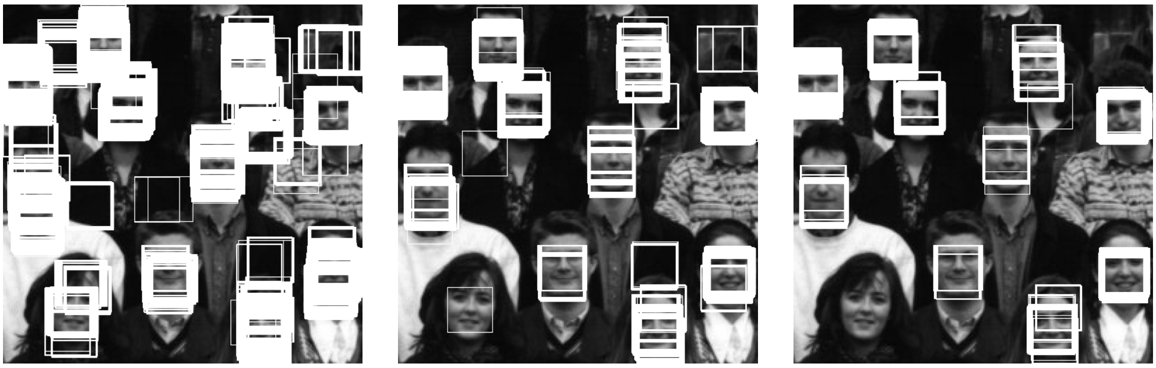
The results shown in (a) and (d) are comparable to those with Manhattan distance. In the work of Ref. [4], the density rule, which remains only high-density areas of the face candidates, was employed to reduce the number of false positives. In this experiment, however, the number of false positives was effectively reduced, without using the density rule, by just increasing the value of p. Furthermore, better performance was obtained when the distance measure using s_i' index was employed. The results with s_i' index were fairly stable with the increase in the value of p (the results with $p = 30$ are almost the same as those of with $p = 10$). In comparison between the results in (c) and (f), only one face was failed to be detected in (f), while two faces were missed in (c). In addition, false positives were effectively reduced in (f). Such an effective elimination of false positives has been achieved by just using the new distance measure without employing advanced algorithms such as the multiple-clue method [3] or the density rule [4]. When the false positives are reduced in the segmentation stage by using the proposed distance measure, the computation load in the verification process using the algorithm proposed in Ref. [6] will be greatly reduced.

3.2 Cephalometric Landmark Identification System

The system developed in Ref. [5] searches the specific anatomical points in cephalometric radiographs. In the system, the minimum distance point is identified as the target point. This time, we searched for Sella, which lies in the center of the pituitary grand. The results of cephalometric landmark identification employing the weighted distance measure using s_i' index are shown in Fig. 4. The minimums of distances occur at two locations in the radiograph, one is at the true Sella and the other at the false Sella. In the case of Manhattan distance, the two minimum values are almost the



(a) s_i index ($p = 0.5$) (b) s_i index ($p = 3$) (c) s_i index ($p = 10$)



(d) s'_i index ($p = 0.5$) (e) s'_i index ($p = 3$) (f) s'_i index ($p = 10$)

Fig. 5. Results of face detection with weighted distance by two indexes, s_i (topside) and s'_i (downside). p is a parameter as in Eq. (1)

same. Actually, the false Sella was detected by the system. In order to eliminate such an error, a macro vision search algorithm in which the search is carried out at two different resolutions was proposed in Ref. [5]. In the case of the new weighted distance the minimum value is clearly larger for the false Sella than for the true Sella. Therefore, the system identified the correct location of Sella without the macro vision search algorithm. We can conclude that the introduction of the proposed weighted distance measure can simplify the procedure of image recognition.

4. CONCLUSION

We have presented a weighed distance measure which uses newly proposed weighting scheme. This scheme calculates the significance factor for each vector element taking its statistics into account. Presented results show that the proposed distance measure is capable of solving various problems encountered in image classification tasks which have been solved by using advanced algorithms.

REFERENCES

[1] H. Lin and A. N. Venetsanopoulos, "A Weighted Minimum Distance Classifier for Pattern Recognition," in *Proc. Canadian Conference on Electrical and Computer Engineering*, Sept. 14-17, 1993, vol. 2, pp. 904–907.

[2] Soonil Kwon and Shrikanth Narayanan, "Speaker Change Detection using A New Weighted Distance Measure," in *Proc. International Conference on Spoken Language Processing*, Denver, United States, Sept. 16-20, 2002, vol. 4, pp. 2537–2540.

[3] Y. Suzuki and T. Shibata, "Multiple-Clue Face Detection Algorithm using Edge-based Feature Vectors," in *Proc. ICASSP 2004*, Montreal, Canada, May 17-21, 2004, vol. 5, pp. 737–740.

[4] Y. Suzuki and T. Shibata, "An Edge-Based Face Detection Algorithm Robust Against Illumination, Focus, and Scale Variations," in *Proc. EUSIPCO 2004*, Vienna, Austria, Sept. 6-10, 2004, pp. 2279–2282.

[5] M. Yagi, and T. Shibata, "An Image Representation Algorithm Compatible With Neural-Associative-Processor-Based Hardware Recognition Systems," *IEEE Trans. Neural Networks*, vol. 14, no. 5, pp. 1141–1161, Sept. 2003.

[6] Y. Suzuki and T. Shibata, "Multiple-Resolution Edge-Based Feature Representations for Robust Face Segmentation and Verification," this conference.