Proc. SPIE's Optical Science and Technology, 49th Annual Meeting, 2 - 6 August 2004 in Denver, Colorado USA, Vol. 5558, 6 p.

Pattern recognition based on rank correlations

Vitaly Kober[a], Mikhael Mozerov[b], Josué Alvarez-Borrego[c]

[a]Departamento de Ciencias de la Computación, División de Física Aplicada, CICESE
Km 107 Carretera Tijuana-Ensenada, Ensenada 22860, B.C., México

[b]Institute for Information Transmission Problems
Bolshoi Karetnii 19, Moscow, 101447 Russia

[c]Departamento de Óptica, División de Física Aplicada, CICESE
Km 107 Carretera Tijuana-Ensenada, Ensenada 22860, B.C., México

ABSTRACT

Adaptive nonlinear filters based on nonparametric Spearman's correlation between ranks of an input scene computed in a moving window and ranks of a target for illumination-invariant pattern recognition are proposed. Several properties of the correlations are investigated. Their performance for detection of noisy objects is compared to the conventional linear correlation in terms of noise robustness and discrimination capability. Computer simulation results for a test image corrupted by mixed additive and impulsive noise are provided and discussed.

Keywords: Rank-order filters, pattern recognition, nonparametric rank correlation.

1. INTRODUCTION

Since the introduction of the matched spatial filters[1] a lot of efforts have been undertaken to improve pattern recognition. The classical matched filter is optimal for the detection of objects in additive Gaussian noise, however, it is not able to discriminate effectively an object of one class and that belonging to other classes. Moreover, it yields a poor performance when an input scene is corrupted by non-Gaussian noise possessing heavy tails in its distribution. Many attempts were undertaken to introduce correlation filters with better performance with respect to various criteria.[2-5] Several performance measures for correlation filters have been proposed and summarized.[2] Some of these measures can be essentially improved using an adaptive approach to the filter design. According to this concept we are interested in a filter with good performance characteristics for a given observed scene, i.e. with a fixed set of patterns or a fixed background to be rejected, rather than to construct a filter with average performance parameters over an ensemble of images. For non-stationary background noise (space-varying data) and unknown illumination of the target, it is preferable to utilize an adaptive processing in a moving window.

It can be shown that a minimization of the mean squared error between an input image and a shifted version of the target under certain assumptions leads to maximizing the linear correlation between the image and the target. From statistical viewpoint, one can claim that the criterion is optimal if the input image corrupted by additive Gaussian noise distributions. However, real data are often corrupted by non-Gaussian noise. In this case other error criteria[6-9] are more robust even for slight deviations from the Gaussian assumption. The standard (Pearson) linear correlation calculations are based on the assumption that signals of an input image and a target are sampled from populations that follow a Gaussian distribution, at least approximately. In order to improve the performance of the linear correlation for pattern recognition with respect to robustness to high-noise situations, we utilize nonparametric (Spearman) correlation[10] instead.

Further author information –
V.K.(correspondence): email: vkober@cicese.mx, telephone: +52-646-1750500; Fax: +52-6-1750593
M.M.:email: mozer@iitp.ru; J.A.-B.: email: josue@cicese.mx

This correlation does not require the Gaussian assumption because it is based on ranking the two variables, and so makes no assumption about the distribution of the values. The Spearman's correlation belongs to rank-order filers.[11-15] These filters have proven to be very effective in removal of additive and impulsive noise. Moreover they exhibit excellent robustness properties.

The paper is organized as follows. In Section 2, we introduce our notations and propose nonlinear adaptive correlations. In Section 3, we provide computer simulation results and compare them in terms of discrimination capability for the proposed filters and noisy test images. Section 4 summarizes our conclusions.

2. RANK CORRELATIONS

The rank filtering is a local adaptive processing of the signal in a moving window. Let us introduce some notations and definitions: $\{s(n,m)\}$ is an input image containing an object to be recognized; (n,m) are the coordinates of pixels, $n=1,2,\ldots N$ and $m=1,2,\ldots M$; $L=N\mathrm{x}M$ is the image matrix size; $\{t(n,m)\}$ is a noiseless target image. A spatial neighborhood for each image pixel can be defined as a set of pixels surrounding the given pixel geometrically. The neighborhood is referred to as the *W*-neighborhood. The size of the neighborhood is referred to as $|W|$ and it is approximately taken as the size of the target. In the case of non-stationary additive noise or cluttered background (space-varying data), it is assumed that the size of the *W*-neighborhood sufficiently small and the signal and noise can be approximately considered stationary over the window area. Let $m_t = \frac{1}{|W|}\sum_{n,m\in W} t(n,m)$ be the average of the target over the *W*-neighborhood, $m(k,l) = \frac{1}{|W|}\sum_{n,m\in W} s(n+k,m+l)$ be the local average of the input image. An important notion in order statistics is the variational row, which is defined as a one-dimensional sequence $\{V(r)\}$ of *K* pixels whose elements are sorted in ascending order with respect to their values: $\{V(r)\colon V(r)\leq V(r+1),\ r=1,2,..K\}$. Here $V(r)$ and $r(V)$ are called the *r*th order statistics and the rank of the value *V*, respectively. Both the rank and any order statistics can be computed from the local histogram of the signal distribution over the *W*-neighborhood centered at each pixel as follows:

$$r(V) = \sum_{q=0}^{V} h(q). \tag{1}$$

All the parameters of rank-order filters are functions of local, or short time, histograms computed over pixels of the spatial neighborhoods. Therefore, the computational complexity of the rank processing depends on calculation of local histograms.

Standard (Pearson) linear correlation at the coordinates (k,l) between an input image and a shifted version of the target can be defined as

$$L(k,l) = \sum_{n,m\in W} t^o(n,m)\, s^o(n+k,m+l) \bigg/ \sqrt{\sum_{n,m\in W} \left|s^o(n+k,m+l)\right|^2 \sum_{n,m\in W} \left|t^o(n,m)\right|^2} \quad , \tag{2}$$

where the sum is taken over the *W*-neighborhood, $t^o(n,m)= t(n,m)-m_t$ and $s^o(n+k, m+l)= s(n+k, m+l)-m(k,l)$ are the centered target and locally centered input image, respectively.

This correlation measures the angular separation of expression vectors for the input image and a shifted version of the target. The Pearson's correlation is one of many possibilities. To use it one must assume that the data in the pairs come from the Gaussian distribution and the data is at least in the category of equal interval data. If these two conditions are not met, a possibility is to use the Spearman's rank correlation. Instead of using the actual data, each the target data point is ranked from 1 to $|W|$ as well as each input image local fragment data point. A nonparametric (distribution-free) rank statistic proposed by Spearman as a measure of the strength of the associations between two variables.[10] It can be defined for pattern recognition as follows:

$$R(k,l)=\frac{\sum\limits_{n,m\in W}\left(r_t(n,m)-\overline{r}_t\right)\left(r_s(n+k,m+l)-\overline{r}_s(k,l)\right)}{\sqrt{\sum\limits_{n,m\in W}\left(r_t(n,m)-\overline{r}_t\right)^2\sum\limits_{n,m\in W}\left(r_s(n+k,m+l)-\overline{r}_s(k,l)\right)^2}}, \quad (3)$$

where all sums are taken over the *W*-neighborhood, $\{r_t(n,m)\}$} are ranks of the target, $\{r_s(n+k,m+l)\}$} are ranks of the input image over the *W*-neighborhood at the coordinates (k,l), $\overline{r}_t$ is the average of the target ranks, . $\overline{r}_s(k,l)$ is the average of the input image ranks over the *W*-neighborhood.

It may be shown that the rank correlation in Eq. (3) after some manipulations is simplified to

$$R(k,l)=1-\frac{6\sum\limits_{n,m\in W}\left(r_t(n,m)-r_s(n+k,m+l)\right)^2}{|W|\left(|W|^2-1\right)}. \quad (4)$$

The rank correlation is a statistical measure of the amount of monotonic relationship between two variables that are arranged in rank order. It can be used to give an *R*-estimate, and is a measure of monotone association that is used when the distribution of the data make Pearson's correlation undesirable or misleading. In the Spearman correlation only the order of the data is important, not the level, therefore extreme variations in expression values have less control over the correlation. Note that ranks possess an attractive feature; that is, they are invariant to any independent monotonic transformations of both an input scene and a target. Therefore the rank correlation is also invariant to any monotonic transformation of an input scene. For instance, a nonuniform illumination of an input image can be described by a monotonic element-wise transformation.

The rank correlation given in Eq. (3) is based on calculations of sums of products. From a statistical point of view, the sum operation might be inadequate in high-noise situations. In particular, the tendency of the sum operation to mimic a single aberrant data value was recognized to be a serious problem with that estimator. In other words, if a data point is not consistent with others (a point corrupted by impulsive noise or an outlier), it will tend to pull the sum in its direction. This is why we propose to incorporate some degree of data editing whereby aberrant data points could be removed from the computation. This leads to robust statistical estimators such as the trimmed sum[11] and its natural extension, the median.[12] The trimmed sum is so called because, rather that summing the entire data set, a few data points are trimmed and the remainder is summed. The points, which are removed, are the most extreme values, both low and high, with a given number of points dropped at each end. In practice, the trimmed sum is computed by sorting the data from low to high and summing the central part of the ordered sequence. The number of data values, which are dropped from the sum, is controlled by the trimming parameters left alpha α_L and right alpha α_R, which assume values between 0 and 0.5. We propose to exclude from the calculation in Eq.(3) $\alpha_L\,|W|$ left ranks and $\alpha_R\,|W|$ right ranks of the target and the input scene. A new trimmed rank correlation is denoted as $R_{\alpha_L,\alpha_R}(k,l)$. When left and right alphas are zeros, the range of summation is over the entire *W*-neighborhood and the trimmed rank correlation is equal to the Spearman's rank correlation.

In next section, with the help of computer simulation we show that the rank and trimmed rank correlations yield much better recognition than the conventional linear correlation when the input scene is distorted by a nonlinear monotonic illumination and noise distribution has heavy tails.

3. COMPUTER SIMULATIONS

In this section, we illustrate the performance of the proposed correlations in terms of discrimination capability (DC) and their sensitivity to additive and impulsive noise in a scene signal. Fig. 1 (a) is an example of the input scene containing a target. The local mean and variance of nonstationary background are varying. The mean value and standard deviation of the entire background are 128 and 25, respectively. The target is imbedded into the background. The test image contains 64x64 pixels. The size of the target area is 23x15. The mean value and standard deviation of the target are 128 and 30, respectively. The W-neighborhood is equal to the region of support of the target. The DC is formally defined as follows:

$$DC = 1 - \frac{\left|C^B\right|^2}{\left|C^O\right|^2}, \tag{5}$$

where C^B is the maximum in the correlation plane over the background area to be rejected, and C^O is the maximum in the correlation plane over the area of object to be recognized. The area of the object to be recognized is determined in the close vicinity of the target location (the size of the area is similar to the size of the target). The background area is complementary to the object area.

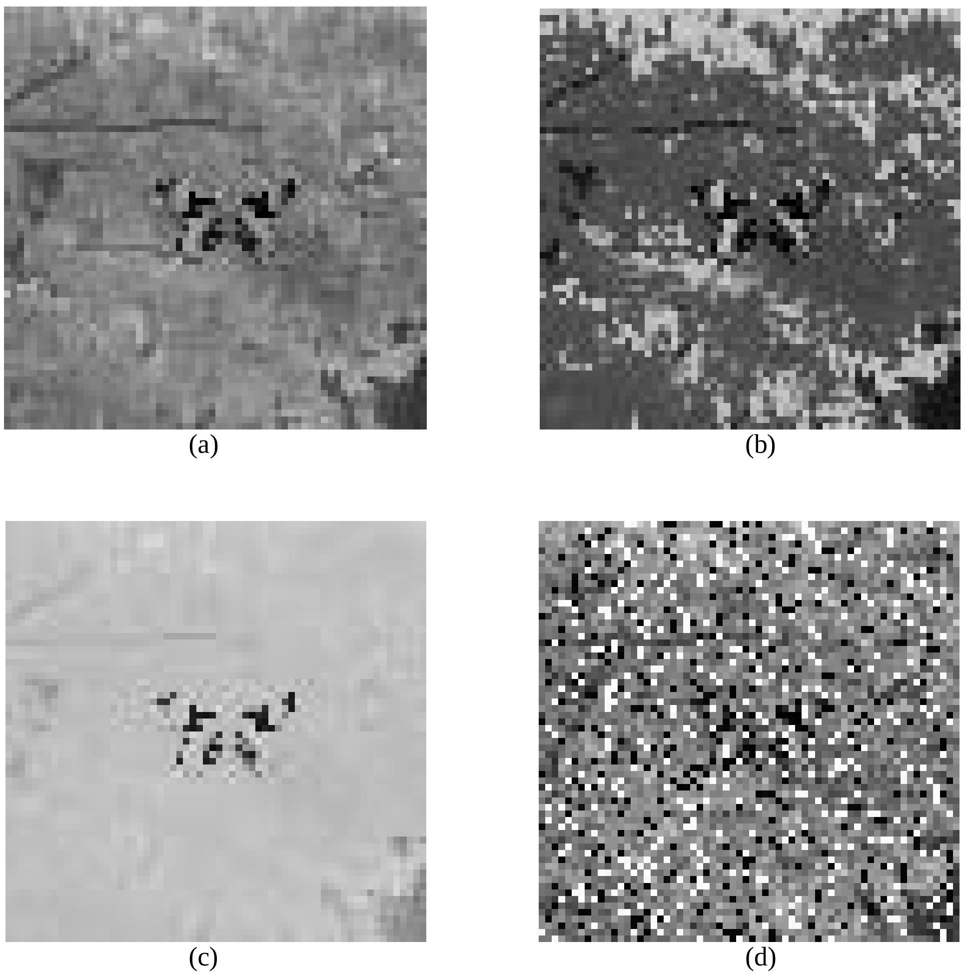

Fig. 1. (a) Test input scene, (b), (c) versions of monotonically transformed test image (d) test image corrupted by additive and impulse noise.

Figs. 1 (b) and (c) show two versions of nonuniform illumination of the input scene obtained with element-wise monotonic transformations. The rank correlation gives the same DC for all images in Fig 1 (a), (b), and (c).

Suppose that the input scene is corrupted by mixed additive and impulsive noise. The additive noise is zero-mean Gaussian distributed with the standard deviations of 5, 10,…,40. The probability of impulse occurring is 0.2, and if it occurs it can be positive or negative with equal probability. Figure 1 (d) shows the input scene corrupted by additive noise with the standard deviation of 20 and degraded due to impulsive noise. For each input noise parameters used here, 30 statistical trials are conducted. The magnitudes of the DC computed over the resulting correlation plane after each statistical trial are averaged. Note that the classical linear correlation fails to recognize the target in the input scene as well in the noisy input scene. The performance of rank correlation filters $R\left(k,l\right)$ and $R_{\alpha_L,\alpha_R}\left(k,l\right)$ in terms of the DC as a function of the standard deviation of additive noise is given in Fig. 2. For the case of the trimmed rank correlation the left and right alphas are chosen to be 0.1. Note that the proposed nonlinear correlations provide a significant improvement in recognition process comparing with the linear correlation filter. Comparing the two rank filters in Fig. 2, we see that the trimmed rank correlation is a more robust filter to input additive and impulsive noise.

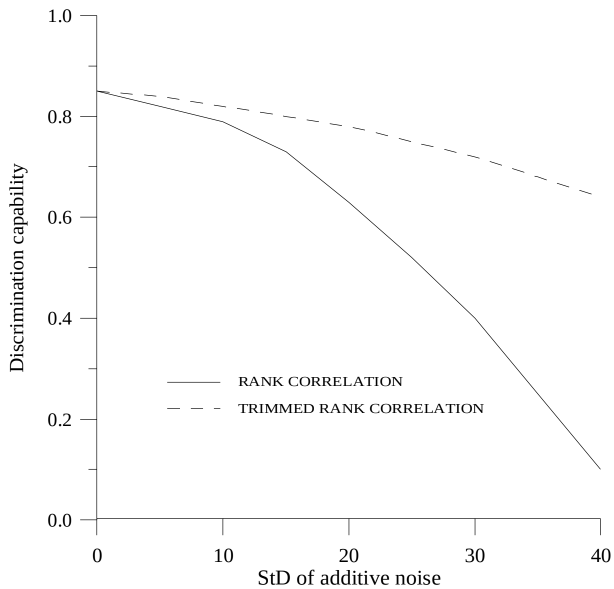

Fig 2. Discrimination capability as a function of the standard deviation of mixed additive noise obtained with rank adaptive correlations.

Next we illustrate the performance of the rank correlations for pattern recognition when an input scene is scaled down. Fig. 3 (a) is an example of the input scene containing three versions of a target. Intensities of objects to be recognized are changed with three different monotonic transformations. The local mean and variance of nonstationary background are varying. The test image contains 256x256 pixels. The size of the target area is 23x15. The W-neighborhood is equal to the region of support of the target. The discrimination capability is calculated with Eq. (5), where C^O is chosen as a minimal maximum in the correlation plane over the areas of object to be recognized. The discrimination capability of the rank correlation is equal to 0.71.

(a) (b)

Fig. 3. (a) Test input scene, (b) test image scaled down by factor of 0.9.

Fig. 3 (b) shows a reduced test image by factor of 0.9. The discrimination capability is reduced to 0.4. Note that in this case the rank correlation is still able to provide a correct recognition.

4. CONCLUSION

Local adaptive nonlinear correlations based on rank statistical estimation were proposed. Their performance for detection of an object corrupted by mixed additive and impulsive noise were compared to the linear correlation in terms of noise robustness and discrimination capability. Computer simulations have clearly illustrated an improvement of pattern recognition when the proposed filters are involved in the recognition process in high-noise situations.

ACKNOWLEDGMENT

The authors acknowledges grant 36077-A from CONACYT.

REFERENCES

1. A. VanderLugt, "Signal detection by complex filters," *IEEE Trans. Inf. Theory*, 1965, vol. IT-10, pp.139-145.
2. V.B.V.K. Kumar and L. Hassebrook, "Performance measures for correlation filters", *Applied Optics*, 1990, vol. 29, pp.2997-3006.
3. L.P. Yaroslavsky, "The theory of optimal methods for localization of objects in pictures," in: *Progress in optics XXXII*, E. Wolf, ed., Elsevier Science Pub. B.V., 1993, pp. 145-201.
4. B. Javidi and J. Wang, "Design of filters to detect a noisy target in nonoverlapping background noise, *Journal OSA (A)*, 1994, vol. 11, pp. 2604-2612.
5. V. Kober and J. Campos, "Accuracy of location measurement of a noisy target in a nonoverlapping background," *Journal OSA (A)*, 1996, vol. 13, pp.1653-1666.
6. E.J. Coyle, "Rank order operators and mean absolute error criterion," *IEEE Trans. Acoust. Speech Signal Process.*, 1988, vol. ASSP-362, pp. 1183-1188.
7. P. Maragos, "Morphological correlation and mean absolute error criteria," *Proc. conf. IEEE Trans. Acoust. Speech Signal Process.*, 1989, Glasgow, pp. 1568-1571.
8. V. Kober, "Robust nonlinear correlations", *Proc. conf. SPIE*, 2003, vol. 5203, pp. 82-87.
9. V. Kober, M. Mozerov, J. Alvarez-Borrego, I.A. Ovseyevich, "Adaptive rank-order correlations", *Pattern Recognition and Image Analysis*, 2004, vol. 14, No. 1, pp. 33-39.
10. R.V. Hogg and A.T. Craig, *Introduction to Mathematical Statistics,* Macmillan, New York, 1995.
11. J.B. Bednar and T.L. Watt, "Alpha-trimmed means and their relationship to median filters," *IEEE Trans. Acoust. Speech Signal Process.*, 1984, vol. ASSP-32, pp. 145-153.
12. J.P. Fitch, E.J. Coyle, and N.C. Gallagher, Jr., "Median filtering by threshold decomposition," *IEEE Trans. Acoust. Speech Signal Process.*, 1984, vol. ASSP-32, pp. 1183-1188.
13. V. Kober, M. Mozerov, J. Alvarez-Borrego, "Nonlinear filters with spatially-connected neighborhoods," *Optical Engineering*, 2001, vol.40, No. 6, pp. 971-983.
14. I. Pitas and A.N. Venetsanopoulos, *Nonlinear digital filters. Principles and applications*, Kluwer Academic Publishers, Boston, 1990.
15. E. D. Dougherty and J. Astola, *Introduction to Nonlinear Image Processing,* Bellingham, WA: SPIE, 1994.