Moment Matrices for Recognition of Spatial Pattern in Noisy Images

A. O. Hero

Abstract

We present a method for detection and classification of a spatial pattern in noise contaminated binary images which is based on performing subspace decomposition on a non-negative definite matrix of higher order moments of the image. We introduce a method which uses normalized power moments or ascending factorial moments as descriptors. While the set of p-th order factorial moments are in one-to-one correspondence with the set of p-th order power moments, the computation of factorial moments is much more numerically stable than the power moments. Indeed, using factorial moments we are able to implement pattern classifiers with twice as many moment descriptors. We illustrate these techniques for word classification in binary document images.

I. Introduction

The problem of classification of patterns in noisy binary images has been a key component to many different areas including: automatic document processing, such as word spotting [1], character recognition [2], database retrieval; automatic target recognition (ATR); and astronomical cartography. We present a pattern classification method which is based on an underlying spatial Poisson point process model and uses higher order factorial moments of the intensity function as pattern discriminants. This extends our work on standard power moment methods presented at ICIP-96 [3]. Some justifications for moments are: 1) they provide a non-parametric pattern description; 2) combinations of moments have been identified with important invariances such as rotation, scale, and translation [4], [5]; 3) sample moments can usually be treated as jointly Gaussian random variables. With our introduction of non-negative definite moment matrices we provide another justification: they can be used to effectively separate signal pattern from noise background via noise subspace processing.

II. Spatial Point Process Representations

Let $W = W(x, y)$ be a binary image indexed over $(x, y) \in \mathcal{A} = \{1, \ldots, n\} \times \{1, \ldots, m\}$. Define the coordinate locations (pixels) $\{(X_i, Y_i)\}_{i=1}^N$ as those pixels at which W is active, i.e., W not equal to zero. A general statistical model for $\{(X_i, Y_i)\}_{i=1}^N$ is a two-dimensional Bernoulli process specified by the probabilities $\rho(x, y) \in [0, 1]$ that pixel (x, y) is active. For a large number of pixels ($nm \gg N$), the Bernoulli process is well approximated by a Poisson point process over the continuous rectangle $A = [1, n] \times [1, m]$. This Poisson point process is completely described by the normalized intensity function $\rho(x, y)$ and the rate constant Λ

$$\lambda(x, y) = \Lambda\rho(x, y), \quad (x, y) \in A$$

where $\Lambda = E[N]$ is the average number of active pixels in the image. The statistical expectation of any function $g(X_i, Y_i)$ is

$$E[g(X_i, Y_i)] = \int_A g(x, y)\rho(x, y)\, dxdy.$$

III. Spatial Moments

The spatial power moment (PM) $\mu_{X,Y}(k, l)$ of (integer) order k and l of positive integer random variables X, Y is defined as:

$$\mu_{X,Y}(k, l) = E[X^k Y^l].$$

Since the bivariate monomials form a basis for the space of all two-dimensional square integrable functions, the set of $2p$ power moments $\{\mu_{X,Y}(k, l)\}_{kl=1}^p$ completely characterizes the image as $p \to \infty$.

Fractional factorial moments come in two varieties: ascending and descending. In this work we use the spatial ascending fractional factorial moment (AFFM) $\alpha_{X,Y}(k, l, s)$ of (integer) order k and l and (real) fraction parameter s which is defined below

$$\alpha_{X,Y}^{kl,s} = E[[X]_s^k\ [Y]_s^l] \tag{1}$$

where we use the Pochammer symbol $[x]_s^k$ to denote the ascending fractional factorial

$$[x]_s^k \quad = \quad x(x + s) \cdots (x + s(k - 1)).$$

By convention we define $[x]_s^0 = 1$. When $s = 1$ and X is a positive integer $[x]_s^k = \frac{(x+k-1)!}{(x-1)!}$ is the standard ascending factorial moment [6].

In [7] we prove that the set of $2p$ PM's $\{\mu_{X,Y}(k, l)\}_{kl=1}^p$ and the set of $2p$ AFFM's $\{\alpha_{X,Y}(k, l, s)\}_{kl=1}^p$ are theoretically equivalent: any of the sets can be expressed mathematically in terms of the other set. However, as will be seen, with finite precision arithmetic, the factorial moment computation is more numerically stable.

A. Power Moment Matrix

We will find it useful to center and scale the variates X, Y to improve numerical stability of the moment matrix computations. Recall that in the present imaging application $X \in \{1, \ldots, m\}$ and $Y \in \{1, \ldots, n\}$. To improve the stability of the PM we replace X and Y by $\tilde{X} = \frac{x-(m+1)/2}{\sigma_p}$ and $\tilde{Y} = \frac{x-(n+1)/2}{\sigma_p}$ where $\sigma_p = (\max(m, n) - 1)/2$. Define the $(2p + 1) \times (2p + 1)$ power moment matrix:

$$\mathbf{P}^{2p} = \Big(\big(\mu_{\tilde{X},\tilde{Y}}(i + j)\big)\Big)_{ij=0,\ldots,p} \tag{2}$$

[1] The author is with the Dept. of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109-2122. This work was supported in part by NSA contract MDA904-95-C-2157 and AFOSR contract FQ8671-9501560.

B. Factorial Moment Matrix

Since ascending factorials moments are not defined for negative variates, construction of an appropriately normalized factorial moment matrix requires more care than for the power moment matrix. First of all we must scale the positive integer variates X and Y differently. Define the variates $\tilde{X} = \frac{x-(m+1)/2}{\sigma_f}$ and $\tilde{Y} = \frac{x-(n+1)/2}{\sigma_f}$ where $\sigma_f = \left[[(\max(m,n)-1)/2]_1^{2p}\right]^{1/(2p)}$. Next, for any real number $T = T^+ - T^-$, we define T^+ as the positive part and T^- as the absolute value of the negative part of T. Define the $(2p+1)\times(2p+1)$ factorial moment matrix $\mathbf{A}^{2p}$:

$$\begin{aligned}\mathbf{A}^{2p} &= \Big(\Big(\alpha_{\tilde{X}^+,\tilde{Y}^+}(i+j,1/\sigma_f) \\ &\quad +(-1)^{i+j}\alpha_{\tilde{X}^-,\tilde{Y}^-}(i+j,1/\sigma_f)\Big)\Big)_{ij=0,\ldots,p}. \end{aligned} \tag{3}$$

IV. Denoising via Eigendecomposition

Here we show that any independent Poisson noise with known moments (up to a constant scale factor) but unknown noise power can be removed from an image via eigendecomposition of the moment matrices. The following important properties of the power moment matrix (2) and the factorial moment matrix (3) are proven in [7]. Both are Hankel matrices which are symmetric non-negative definite. Both matrices have all entries in the interval $[-1,1]$ and are sparse for any intensity which is symmetric about the center point $(\frac{m+1}{2},\frac{n+1}{2})$ of the image (Fig. 1). Finally, by construction the $(1,1)$ entry of these two matrices is equal to 1.

Let a signal-independent binary noise having intensity $\lambda_n(x,y) = \Lambda_n\rho_n(x,y)$ be added (modulo 2) to the image. Denote the noiseless image (signal pattern) intensity by $\lambda_s(x,y) = \Lambda_s\rho_s(x,y)$. Then the overall image intensity $\lambda(x,y) = \Lambda\rho(x,y)$ will be the sum

$$\begin{aligned}\lambda(x,y) &= \lambda_s(x,y)+\lambda_n(x,y) \\ &= \Lambda\left[\beta\rho_s(x,y)+(1-\beta)\rho_n(x,y)\right]\end{aligned}$$

where $\beta = \Lambda_s/(\Lambda_s+\Lambda_n) \in [0,1]$ is a monotone function of signal-to-noise ratio Λ_s/Λ_n. The above representation implies that for any function $g(X,Y)$ of the coordinate process (X,Y) we have the decomposition:

$$\begin{aligned}E[g(X,Y)] &= \beta E[g(X,Y)|\text{signal alone}] \\ &\quad +(1-\beta)E[g(X,Y)|\text{noise alone}]\end{aligned}$$

Let $\mathbf{M}$, $\mathbf{M}_s$, and $\mathbf{M}_n$ denote the signal plus noise, signal alone, and noise alone moment matrices ($\mathbf{M}$ denotes either power moment or factorial moment matrix). We have the important result:

$$\mathbf{M} = \beta\mathbf{M}_s + (1-\beta)\mathbf{M}_n$$

Now let $\mathbf{K}$ be the known (normalized) Cholesky factor of $\mathbf{M}_n$, i.e. $\mathbf{M}_n = \sigma^2\mathbf{K}\mathbf{K}^T$, and define $\tilde{\mathbf{M}} = \mathbf{K}^{-1}\mathbf{M}\mathbf{K}^{-T}$. Then we have:

$$\tilde{\mathbf{M}} = \beta\tilde{\mathbf{M}}_s + (1-\beta)\sigma^2\mathbf{I} \tag{4}$$

where $\mathbf{I}$ is the $(2p+1)\times(2p+1)$ identity matrix, and $\tilde{\mathbf{M}}_s = \mathbf{K}^{-1}\mathbf{M}\mathbf{K}^{-T}$ is the **whitened moment matrix** of the signal pattern.

Except for the presence of the signal-to-noise ratio dependent scalar β, (4) is the standard additive decomposition of a "measurement covariance" matrix $\tilde{\mathbf{M}}$ into unknown "signal covariance" $\tilde{\mathbf{M}}_s$ and "white noise covariance" $\sigma^2\mathbf{I}$. While the latter is full rank with $2p+1$ constant eigenvalues $\{\sigma^2,\ldots,\sigma^2\}$, the eigenvalues of the signal matrix $\tilde{\mathbf{M}}_s$ are non-negative and typically fall off rapidly to zero. This permits exact recovery of $\mathbf{M}_s$ via application of the eigendecomposition: $\tilde{\mathbf{M}} = \sum_{i=1}^{2p+1}\gamma_i\underline{\xi}_i\underline{\xi}_i^T$ where γ_i and $\underline{\xi}_i$ are eigenvalues (rank ordered) and eigenvectors.

To wit, since the eigenvectors of $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{M}}_s$ are identical, from (4) we have

$$\tilde{\mathbf{M}} = \sum_{i=1}^{q}\left[\beta\gamma_i^s+(1-\beta)\sigma^2\right]\underline{\xi}_i\underline{\xi}_i^T + (1-\beta)\sigma^2\sum_{i=q+1}^{2p+1}\underline{\xi}_i\underline{\xi}_i^T$$

where $\{\gamma_i^s\}_{i=1}^q$ are the non-zero eigenvalues of $\tilde{\mathbf{M}}_s$. Thus only the q largest eigenvalues $\gamma_i = \beta\gamma_i^s+(1-\beta)\sigma^2$ of $\tilde{\mathbf{M}}$ are related to the signal pattern and the rest are pure noise eigenvalues $\{(1-\beta)\sigma^2,\ldots,(1-\beta)\sigma^2\}$. In particular, at least if we know β and σ^2 *a priori*, then $\tilde{\mathbf{M}}_s$ can be exactly recovered from the eigedecomposition of $\tilde{\mathbf{M}}$ via

$$\tilde{\mathbf{M}}_s = \frac{1}{\beta}\sum_{i=1}^{q}\left[\gamma_i-(1-\beta)\sigma^2\right]\underline{\xi}_i\underline{\xi}_i^T.$$

However, (surprise!) it turns out that prior knowledge of β and σ^2 is not required. Indeed, since the $(1,1)$ element of $\mathbf{M}_s$ is equal to 1, and the smallest $2p-q+1$ eigenvalues of $\tilde{\mathbf{M}}$ are equal to $(1-\beta)\sigma^2$, we can determine β and $(1-\beta)\sigma^2$ from the two simultaneous linear equations:

$$\begin{aligned}\beta &= \sum_{i=1}^{q}\left[\gamma_i-(1-\beta)\sigma^2\right]|\underline{\xi}_{i1}|^2. \\ (1-\beta)\sigma^2 &= \frac{1}{2p-q+1}\sum_{i=q+1}^{2p+1}\gamma_i\end{aligned}$$

V. A Wordspotting Example

We obtained an electronic copy (ascii) of Thomas Payne's Common Sense, segmented the ascii text, and generated postscript versions of each word in various font sizes, pitches, and font types. Next high levels of additive salt and pepper noise were added to the bitmaps of each word image and the power and factorial moments matrices of varying orders were computed using Matlab. Finally eigendecompositions were computed (signal subspace dimension q was determined by a standard empirical threshold rule) and the denoising procedure discussed above was implemented to recover the original noiseless moment matrix. A representative example is shown in (Fig. 2)-(Fig. 3) for the power moment method. It was found that almost twice as many factorial moments could be included in the analysis without running into run time errors due to numerical roundoff (Fig. 4). Note the distinct features of the recovered whitened matrices for each of the two words ("will" and "war"). Our results suggest that the whitened moment matrices may themselves be an excellent feature space for robust discrimination between words.

References

[1] F. Chen, L. Wilcox, and D. Bloomberg, "Word spotting in scanned images using hidden Markov modeling," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, volume V, pp. 1–4, Minneapolis, MN, 1993.

[2] S. Kahan, T. Pavlidis, and H. Baird, "On the recognition of printed characters of any font and size," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 9, no. 2, pp. 274–288, Mar, 1987.

[3] J. O'Neill, A. O. Hero, and W. J. Williams, "Word spotting via spatial point processes," in *IEEE Int. Conf. on Image Processing*, volume 2, pp. 215–219, Laussane, Switzerland, Sept. 1996.

[4] M. Hu, "Pattern recognition by moment invariants," *Proc. of the Institute of Radio Engineers (IRE)*, p. 1428, 1961.

[5] Y. Li, "Reforming the theory of invariant moments for pattern recognition," *Pattern Recognition*, no. 7, pp. 723–730, 1992.

[6] N. L. Johnson, S. Kotz, and A. W. Kemp, *Univariate discrete distributions (2nd Ed)*, Wiley, New York, 1992.

[7] A. O. Hero, "Moment matrix methods for robust pattern recognition," Technical report, Comm. and Sig. Proc. Lab. (CSPL), Dept. EECS, University of Michigan, Ann Arbor, Jan., 1997.

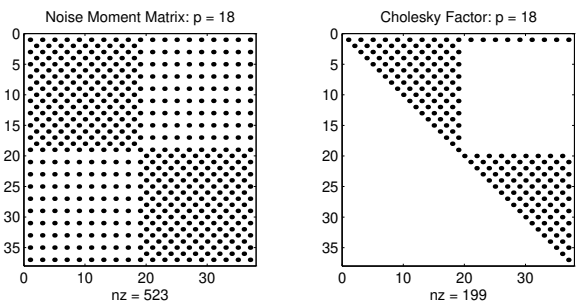

Fig. 1. The noise alone factorial moment matrix and its Cholesky factor have sparse structure which can be exploited to reduce roundoff error and explore structure of moment invariants.

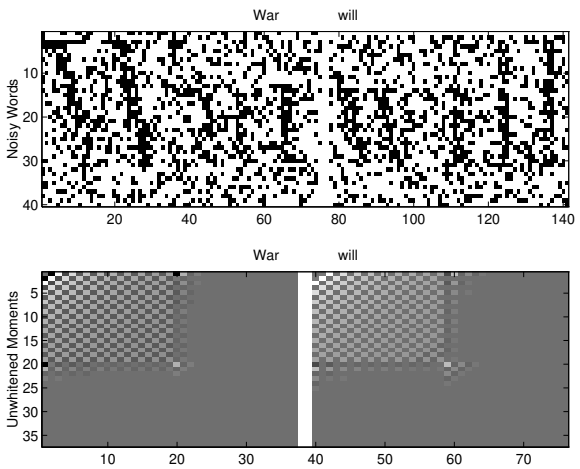

Fig. 2. Top: the words *will* and *War* in helvetica font and severe additive noise. Bottom: the unwhitened power moment matrices empirically estimated from the data strongly resemble noise alone moment matrices and no distinctive features are evident between words.

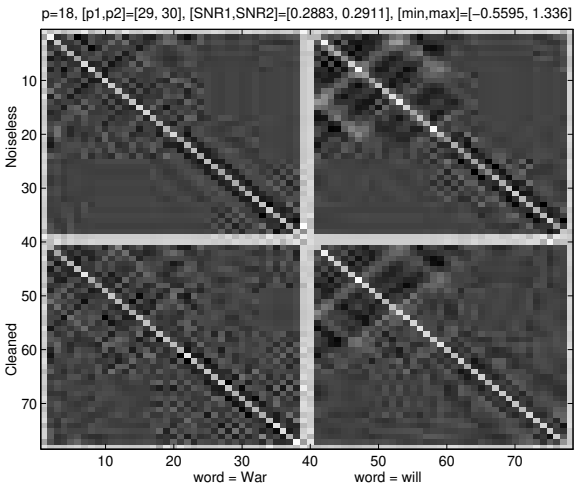

Fig. 3. Top: Noiseless whitened power moment matrices for words *will* and *War* for $p = 18$. Bottom: denoised empirical moment matrix from noisy vesrions of words shown in Fig 2. Note that distinctive features of noiseless whitened power moments are recovered after denoising.

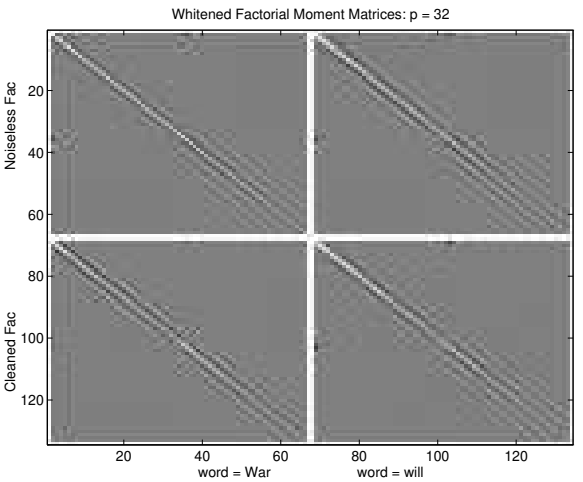

Fig. 4. Same as Fig. 3 except that $p = 32$ factorial moment matrix is shown.