Change Point Detection in a Sequence of Recognized Images

Nicholas A. Nechval, Konstantin N. Nechval & Edgars K. Vasermanis

Department of Applied Mathematics, Aviation University of Riga
Lomonosov Street 1, LV-1019 Riga, Latvia
fax: +371 7 100660, e-mail: nechval@junik.lv

Abstract

In the present paper, we develop a new efficient approach to detection of abrupt changes in a sequence of recognized digital images, which can be applied to a wide variety of practical problems of digital image processing. The principal idea of this approach consists in transforming an original sample of the observed data to a set of one-dimensional statistics and measuring a distance from homogeneity. Essentially, there are two problems associated with change point detection: detecting the change and making inferences about the change point. For solving these problems, a non-parametric technique is proposed. The test for testing the null hypothesis of “no change” against the alternative of “change” is based on a version of the Waerden statistic. Estimating the change point is based on a version of the Mann-Whitney statistic. The suggested approach represents a sort of “retrospective” approach when (at each stage) one looks at a fixed sample of the statistical data and attempts to determine whether and where a change has occurred in this fixed sample of the foregoing data. The main advantage of the proposed approach lies in its relative simplicity and the ease with which it can be applied.

Keywords: digital image, sequence, change point, detection, test

1. Introduction

In many practical problems concerned with digital image processing one is faced with the task of deciding if the variability within a sequence of digital images is uniform, or if not, which segments of a sequence exhibit the uniform variability. This type of problems arises, when the data relate to several processes, to the same process at different times, or to the same processes from different sources. Its applications appear in many fields such as computer vision systems, quality control, reliability theory, seismology, speech processing, automatic analysis of biomedical signals, econometric modelling, ecological modelling, remote sensing, regression analysis and tracking problems.

2. Data Model

Consider a digital image representing a set of p picture elements (pixels) in some object. Let $\mathbf{s}=(s_1, \ldots ,s_p)'$ be the signal associated with this image, where s_i is a gray-level value at the ith pixel, $i\in\{1, \ldots ,p\}$. It is assumed that the image is distorted by a noise process. The problem of detecting the unknown deterministic signal **s** in the presence of a noise process, which is incompletely specified, can be viewed as a binary hypothesis testing problem. The decision is based on a sample of observation vectors $\mathbf{x}_i = (x_{i1}, \ldots ,x_{ip})'$, $i = 1(1)n$, each of which is composed of noise $\mathbf{w}_i = (w_{i1}, \ldots ,w_{ip})'$ under the hypothesis H_0 and a signal $\mathbf{s} = (s_1, \ldots ,s_p)'$ added to noise $\mathbf{w}_i$ under the alternative H_1, where $n > p$. The two hypotheses which the adaptive detector must distinguish are given by

$$H_0: \mathbf{X} = \mathbf{W} \qquad \text{(noise alone)}, \tag{1}$$

$$H_1: \mathbf{X} = \mathbf{W} + \mathbf{cs'} \qquad \text{(signal present)}, \tag{2}$$

where

$$\mathbf{X} = (\mathbf{x}_1, \ldots ,\mathbf{x}_n)', \tag{3}$$

$$\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_n)', \tag{4}$$

are n × p random matrices, and

$$\mathbf{c} = (1, \dots, 1)' \tag{5}$$

is a column vector of n units. It is assumed that $\mathbf{w}_i$, i = 1(1)n, are independent and normally distributed with zero-means and common unknown covariance matrix (positive definite) **Q**, i.e.

$$\mathbf{w}_i \sim N_p(\mathbf{0},\mathbf{Q}), \qquad \forall i = 1(1)n. \tag{6}$$

Thus, for fixed n, the problem is to construct a test which consists of testing the null hypothesis

$$H_0: \mathbf{x}_i \sim N_p(\mathbf{0},\mathbf{Q}), \qquad \forall i = 1(1)n, \tag{7}$$

versus the alternative

$$H_1: \mathbf{x}_i \sim N_p(\mathbf{s},\mathbf{Q}), \qquad \forall i = 1(1)n, \tag{8}$$

where the parameters **s** and **Q** are unknown. One of the possible statistics for testing H_0 versus H_1 is given by the generalized maximum likelihood ratio (GMLR)

$$LR = \max_{\theta\in\Theta_1} L_{H_1}(\mathbf{X};\theta) \Big/ \max_{\theta\in\Theta_0} L_{H_0}(\mathbf{X};\theta), \tag{9}$$

where $\theta=(\mathbf{s},\mathbf{Q})$, $\Theta_0 = \{(\mathbf{s},\mathbf{Q}): \mathbf{s}=\mathbf{0}, \mathbf{Q}\in \mathcal{M}_p\}$, $\Theta_1 = \Theta - \Theta_0$, $\Theta = \{(\mathbf{s},\mathbf{Q}): \mathbf{s}\in \mathbf{R}^p, \mathbf{Q}\in \mathcal{M}_p\}$, $\mathcal{M}_p$ denotes the set of p × p positive definite matrices. Under H_0, the joint likelihood for **X** based on (7) is

$$L_{H_0}(\mathbf{X};\theta) = (2\pi)^{-np/2}|\mathbf{Q}|^{-n/2} \times \exp\left(-\sum_{i=1}^{n} \mathbf{x}_i'\mathbf{Q}^{-1}\mathbf{x}_i / 2\right). \tag{10}$$

Under H_1, the joint likelihood for **X** based on (8) is

$$L_{H_1}(\mathbf{X};\theta) = (2\pi)^{-np/2}|\mathbf{Q}|^{-n/2} \times \exp\left(-\sum_{i=1}^{n} (\mathbf{x}_i-\mathbf{s})'\mathbf{Q}^{-1}(\mathbf{x}_i-\mathbf{s})/2\right). \tag{11}$$

It can be shown that

$$LR = |\hat{\mathbf{Q}}_0|^{n/2}|\hat{\mathbf{Q}}_1|^{-n/2}, \tag{12}$$

where

$$\hat{\mathbf{Q}}_0 = \mathbf{X}'\mathbf{X}/n, \tag{13}$$

$$\hat{\mathbf{Q}}_1 = (\mathbf{X}'-\hat{\mathbf{s}}\mathbf{c}')(\mathbf{X}'-\hat{\mathbf{s}}\mathbf{c}')'/n, \tag{14}$$

and

$$\hat{\mathbf{s}} = \mathbf{X}'\mathbf{c}/n \tag{15}$$

are the well-known maximum likelihood estimators of the unknown parameters **Q** and **s** under the hypotheses H_0 and H_1, respectively. It can be shown, after some algebra, that (12) is equivalent finally to the statistic

$$v = \mathbf{T}_1'\mathbf{T}_2^{-1}\mathbf{T}_1 / n, \tag{16}$$

where $\mathbf{T}_1 = \mathbf{X}'\mathbf{c}$, $\mathbf{T}_2 = \mathbf{X}'\mathbf{X}$. It is known that $(\mathbf{T}_1,\mathbf{T}_2)$ is a complete sufficient statistic for the parameter $\theta = (\mathbf{s},\mathbf{Q})$. Thus, the problem has been reduced to consideration of the sufficient statistic $(\mathbf{T}_1,\mathbf{T}_2)$. It can be shown that under H_0, the result (16) is a **Q**-free statistic, which has the property that its distribution does not depend on the actual covariance matrix **Q**. This is given by the following theorem.

Theorem 1. Under H_1, the statistic v is subject to a noncentral beta-distribution with the probability density function

$$f_{H_1}(v;n,q) = \left[B\left(\frac{p}{2},\frac{n-p}{2}\right)\right]^{-1} v^{\left(\frac{p}{2}\right)-1} \times (1-v)^{\left(\frac{n-p}{2}\right)-1} e^{-\left(\frac{q}{2}\right)} {}_1F_1\left(\frac{n}{2};\frac{p}{2};\frac{qv}{2}\right), \quad 0 < v < 1, \tag{17}$$

where ${}_1F_1(a;b;x)$ is the confluent hypergeometric function, $q = n(\mathbf{s}'\mathbf{Q}^{-1}\mathbf{s})$ is a noncentrality parameter representing the generalized signal-to-noise ratio (GSNR). Under H_0, when q=0, (17) reduces to a standard beta-function density of the form

$$f_{H_0}(v;n) = \left[B\left(\frac{p}{2},\frac{n-p}{2}\right)\right]^{-1} \times v^{\left(\frac{p}{2}\right)-1}(1-v)^{\left(\frac{n-p}{2}\right)-1}, \quad 0 < v < 1. \tag{18}$$

Proof. This is given by Nechval (1992, 1997) and so it is omitted here. □

The GMLR test of H_0 versus H_1, based on v, is given by

$$v \begin{cases} \geq h, & \text{then } H_1 \\ < h, & \text{then } H_0 \end{cases}. \tag{19}$$

For fixed n, in terms of the probability density function (18), tables of the central beta-distribution permit one to choose a threshold h of the test in order to achieve the desired test size (false alarm probability P_{FA}). Furthermore, once h is chosen, tables of the noncentral beta-distribution permit one to evaluate, in terms of the probability density function (17), the power (detection probability P_D) of the test. The following theorem shows that the test (19) is uniformly most powerful invariant for a natural group of transformations on the space of observations.

Theorem 2. For testing the hypothesis H_0 (7) versus the alternative H_1 (8), the test given by (19) is uniformly most powerful invariant (UMPI).

Proof. The proof is similar to that of Nechval (1997) and so it is omitted here. □

3. Problem Statement

Consider, under H_1, a sequence of random variables $v_1, v_2, \ldots, v_T$ ordered in time, then the sequence is said to have a change point at τ if v_t for $t=1, \ldots, \tau$ have a common distribution function $F_{H_1}(v; n, q_1)$ and v_t for $t=\tau+1, \ldots, T$ have a common distribution function $F_{H_1}(v; n, q_2)$, where $q_1 \neq q_2$. The parameters q_1 and q_2 are unknown. We consider the problem of testing the null hypothesis of "no-change", ${}^{C}H_0$: $\tau=T$, against the alternative of "change", ${}^{C}H_1$: $1\le\tau<T$, using non-parametric statistics.

In formal terms, the change point problem, considered in this paper, is to decide (at each stage T) between the two hypotheses:

$$ {}^{C}H_0 : v_t \sim F_{H_1}(v; n, q_1), \qquad t = 1(1)T, \tag{20}$$

and

$$ {}^{C}H_1 : v_t \sim F_{H_1}(v; n, q_1), \qquad t = 1(1)\tau,$$

$$ \sim F_{H_1}(v; n, q_2), \qquad t = \tau + 1(1)T, \tag{21}$$

with the prescribed false alarm probability α. Thus, the problem is to determine whether or not a change point exists in a sequence of random variables v_t, $t=1(1)T$. In other words, we consider the problem of testing the null hypothesis of "no-change", ${}^{C}H_0$: $\tau=T$, against the alternative of "change", ${}^{C}H_1$: $1\le\tau<T$, using a non-parametric statistic. If the hypothesis ${}^{C}H_1$ is accepted, a stopping rule S is determined by S =T. The proposed test procedure prescribes stopping as soon as the non-parametric test statistic for detection of a change having occurred exceeds a threshold corresponding to the significance level α.

4. Estimating the Point of a Probable Change

Let us now examine the Mann-Whitney statistic for testing if two samples ($v_1, v_2, \ldots, v_t$ and $v_{t+1}, \ldots, v_T$) come from the same population. The statistic $U_{t,T}$ is defined as

$$ U_{t,T} = \sum_{i=1}^{t} \sum_{j=t+1}^{T} D_{ij}, \tag{22}$$

where $D_{ij}=\mathrm{sgn}(v_i-v_j)$, $\mathrm{sgn}(x)=1$ if $x>0$, 0 if $x=0$, -1 if $x<0$. To use the above statistic to solve the change point problem, we let t vary such that $1\le t<T$. Then we introduce the following statistics:

$$ G_T = \max_{1\le t<T} \left|U_{t,T}\right|, \tag{23}$$

$$ G_T^{+} = \max_{1\le t<T} U_{t,T}, \tag{24}$$

$$ G_T^{-} = -\min_{1\le t<T} U_{t,T}. \tag{25}$$

We refer to G_T, G_T^{+}, and G_T^{-} as the change point statistics. It is easy to see that

$$ G_T = \max\left(G_T^{+}, G_T^{-}\right). \tag{26}$$

An estimation of a probable change point in a sequence of observations is based on this statistic. Unfortunately, a probability distribution of G_T is unknown. Therefore, for testing the null hypothesis ${}^{C}H_0$ against the alternative ${}^{C}H_1$ we use the Waerden statistic (Waerden, 1953).

The above statistic $U_{t,T}$ is used for estimating the point of a probable change. We let t vary such that $1 \le t < T$. The formula (22) can be computationally expensive. It is desirable to use a method that is computationally feasible. We present an alternative to equation (22). Note that

$$U_{t,T} = \sum_{i=1}^{t-1} \sum_{j=t+1}^{T} D_{ij} + \sum_{j=t+1}^{T} D_{tj}. \qquad (27)$$

Now

$$\sum_{i=1}^{t-1} \sum_{j=t+1}^{T} D_{ij} = \sum_{i=1}^{t-1} \sum_{j=t}^{T} D_{ij} - \sum_{i=1}^{t-1} D_{it}. \qquad (28)$$

Substituting (28) into (27), we have

$$U_{t,T} = \sum_{i=1}^{t-1} \sum_{j=t}^{T} D_{ij} - \sum_{i=1}^{t-1} D_{it} + \sum_{j=t+1}^{T} D_{tj}. \qquad (29)$$

If we observe that

$$U_{t-1,T} = \sum_{i=1}^{t-1} \sum_{j=t}^{T} D_{ij}, \qquad (30)$$

$D_{tt}=0$ by definition,

$$-\sum_{i=1}^{t-1} D_{it} = \sum_{i=1}^{t-1} D_{ti} \qquad (31)$$

by symmetry of the sgn function, we arrive at our final recursion formula

$$U_{t,T} = U_{t-1,T} + \sum_{i=1}^{T} D_{ti}. \qquad (32)$$

Then an estimate of the point of a probable change is given by

$$t^* = \arg \max_{1 \le t < T} |U_{t,T}|. \qquad (33)$$

5. Test Statistic for Recognition of a Change

Let $v_1, \ldots, v_{t^*}$ and $v_{t^*+1}, \ldots, v_T$ be independent observed variables. Suppose $r_1, \ldots, r_{t^*}$ are the ranks of the t* observations $v_1, \ldots, v_{t^*}$ in the complete sample of T observations. Let Φ be the (cumulative) distribution function of the standard normal distribution and $\Psi=\Phi^{-1}$ the inverse function. Put

$$b_r = \Psi\left(\frac{r}{T+1}\right), \quad r = 1, \ldots, T. \qquad (34)$$

The hypothesis ${}^{C}H_0$ to be tested is: the v's from $v_1, \ldots, v_{t^*}$ have the same distribution as the v's from $v_{t^*+1}, \ldots, v_T$. The test statistic is

$$Z = \sum_{i=1}^{t^*} b_{r_i}. \qquad (35)$$

If Z exceeds a limit z_α depending on the level α, the hypothesis ${}^{C}H_0$ is rejected. The two-sided test on the level 2α rejects when the absolute value $|Z|$ exceeds the same limit z_α.

Put $b_1=-b$ and $b_T=+b$. Then the probability distribution function F(z) of the statistic Z, that is, the probability of Z < z, is given approximately by

$$F(z) = (1/4)\left[2\Phi\left(\frac{z}{\sigma}\right)+\Phi\left(\frac{z-b}{\sigma}\right)+\Phi\left(\frac{z+b}{\sigma}\right)\right], \quad (36)$$

where

$$\sigma=(t^*-1)\left[\frac{1}{(n-2)(n-3)}\sum_{r=2}^{T-1}b_r^2\right]^{1/2}. \quad (37)$$

The null hypothesis of no change is rejected if the absolute value $|Z|$ exceeds a limit depending on the level α, and we declare that $\tau=t^*$ (see (33)). In other words,

$$\alpha\begin{cases}<\alpha_Z, & \text{then} \quad {}^{C}H_0 \quad \text{(no change)},\\ \geq\alpha_Z, & \text{then} \quad {}^{C}H_1 \quad \text{(change)},\end{cases} \quad (38)$$

where α is a specified significance level,

$$\alpha_Z=1-F(|Z|)+F(-|Z|). \quad (39)$$

Under the hypothesis ${}^{C}H_1$, the maximum likelihood estimates of the unknown parameters $q_1$and q_2 can be found as

$$\hat{q}_1=\arg\max_q\prod_{t=1}^{\tau}f_{H_1}(v_t;n,q) \quad (40)$$

and

$$\hat{q}_2=\arg\max_q\prod_{t=\tau+1}^{T}f_{H_1}(v_t;n,q). \quad (41)$$

6. Conclusion

The results of computer simulation confirm the validity of the theoretical predictions of performance of the suggested test.

Acknowledgments

This research was supported in part by Grant No. 01.0031 and Grant No. 02.0918 from the Latvian Council of Sciences and the National Institute of Mathematics and Informatics of Latvia.

References

Nechval, N.A. (1992). Radar CFAR thresholding in clutter under detection of airborne birds. In: *Proceedings of the 21st Meeting of Bird Strike Committee Europe.* Jerusalem, Israel, pp.127-140.

Nechval, N.A. (1997). Adaptive CFAR tests for detection of a signal in noise and deflection criterion. In: *Digital Signal Processing for Communications Systems* (edited by T. Wysocki, H. Razavi and B. Honary). Kluwer Academic Publishers, Boston/Dordrecht/London, pp.177-186.

Van der Waerden, B.L. (1953). Ein neuer test für das problem der zwei stichproben. *Math. Annalen*, Vol. 126, pp. 93-107.