Integration of Bayes Detection and Target Tracking in Real Clutter Image Sequences

Marcelo G. S. Bruno(1) and José M. F. Moura(2)

(1) Elect. Eng. Depart., University of São Paulo, São Paulo, SP 05508-090, Brazil

ph: (55-11) 3818-5290; fax: (55-11) 3818-5128; email:bruno@lcs.poli.usp.br

(2) Depart. Elect. and Comp. Eng., Carnegie Mellon University

5000 Forbes Avenue, Pittsburgh PA, 15213-3890

ph: (412) 268-6341; fax: (412) 268-3890; email:moura@ece.cmu.edu

Abstract— **We present in this paper an optimal Bayesian algorithm for integrated, multiframe detection and tracking of dim targets that move randomly in spatially correlated, cluttered image sequences. The algorithm consists of a multiframe minimum probability of error Bayes detector integrated with a multiframe maximum a posteriori (MAP) position estimator. The design of the detector/tracker incorporates the models for target signature, target motion, and clutter; it uses recursive spatio-temporal processing across all available frames to make detection decisions and to generate position estimates. A simulation with an artificial target template added to real clutter background shows that the proposed algorithm outperforms the association of a standard single frame image correlator and a linearized Kalman-Bucy filter in a scenario of heavy clutter.**

I. Introduction

We discuss in this paper a new algorithm for integrated detection and tracking of dim targets in sequences of finite resolution two-dimensional (2D) cluttered images. Common approaches to the problem, see [1], are based on a suboptimal decoupling of detection and tracking, i.e., the measurements of interest to the tracker are not the raw sensor images, but the output of a preliminary detection subsystem. The detection stage involves the thresholding of the raw data, usually one single sensor frame. After further preprocessing, validated detections provide measurements that, for targets declared present, are treated as noise-corrupted observations of the unknown true target state (for example, spatial position). A linearized dynamic model is associated to the state of each target, and a tracking filter, usually a variation on the Kalman-Bucy filter, combines the validated measurements from the detection stage with the dynamic model to estimate the target's state. Although the suboptimal combination of single frame detection and Kalman-Bucy filter tracking may perform well in scenarios of high target-to-clutter ratio, its performance tends to deteriorate in a situation where it is necessary to detect and track dim targets in heavily cluttered environments.

Rather than decoupling detection and tracking as in [1], or considering spatio-temporal detection-only (no tracking) as in [2], [3], [4], we present the optimal, multiframe, Bayes detector/tracker that processes directly the sensor images and integrates detection and tracking into a unified framework. The optimal Bayesian algorithm that we describe takes advantage of all prior information on the clutter, target signature, and target motion models, and allows multiframe detection <u>and</u> tracking with recursive spatio-temporal processing across all observed sensor frames.

This paper is divided into six sections. Section I is this introduction. Section II briefly reviews the models for target and clutter that underly our integrated approach to detection and tracking. In section III, we describe the optimal Bayes detector/tracker. Section IV examines the problem of estimating the parameters of the clutter model from the data using an approximate maximum likelihood algorithm. We present in section V an application of the proposed Bayes detector/tracker to infrared airbone laser radar (IRAR) [5] data. The Bayes tracker is shown to outperform the suboptimal association of a standard 2D image correlator with a linearized Kalman-Bucy filter. Finally, section VI summarizes the contributions of the paper.

II. The Model

<u>Clutter</u> After removing the spatially-varying local mean, we describe the background clutter in each frame using a 2D, zero-mean, noncausal, spatially homogeneous, *correlated* Gauss-Markov random field (GMrf) model [6]. For a first order model, the random component of the clutter intensity at pixel (i, j) in the nth frame is modeled by the minimum mean squared error (MMSE) representation

$$V_n(i,j) = \beta_h\,[V_n(i,j-1) + V_n(i,j+1)] + \beta_v\,[V_n(i-1,j) + V_n(i+1,j)] + U_n(i,j)$$

where $1 \le i \le L$, $1 \le j \le M$, and U_n is the prediction error such that

$$E\,[V_n(i,\,j)U_n(k,l)] = \sigma_u^2 \delta(i-k,\,j-l)\,. \tag{1}$$

In (1), $E\,[.]$ stands for expected value (or ensemble average) and δ is the 2D Dirac delta function.

Target Signature We assume a single target scenario where the target of interest is a rigid body whose clutter-free image is contained inside a 2D rectangular region of size $(r_i+r_s+1)\times(l_i+l_s+1)$. In this notation, r_i and r_s denote the maximum vertical pixel distances in the target image when we move away, respectively up and down, from the target centroid. Analogously, l_i and l_s are the maximum horizontal pixel distances in the target image when we move away, respectively left and right, from the target centroid.

The target signature is described by a set of signature coefficients

$$a^n(k,\,l) = c^n(k,\,l)\,\phi^n(k,\,l) \tag{2}$$

for $-r_i \le k \le r_s$ and $-l_i \le l \le l_s$. In (2), $c^n(k,\,l) \in \mathcal{B} = \{0,\,1\}$ is a binary coefficient that defines the target's shape, whereas $\phi^n(k,\,l) \in \Re$ is a real coefficient that specifies the target's pixel intensities. In this paper, we consider the case when the signature coefficients $a^n(k,\,l)$ are deterministic and known to the detector/tracker at each frame. For a treatment of targets with unknown, random pixel intensities, see [7].

Target Motion Due to the sensor's finite resolution, the surveillance space is discretized by a uniform 2D finite lattice. To model situations when targets move in and out of the sensor grid, we define the *centroid lattice* $\widehat{\mathcal{L}} = \{(i,j)\colon -r_s+1 \le i \le L+r_i,\ -l_s+1 \le j \le M+l_i\}$, where L and M are the number of resolution cells in each dimension. Let $\overline{\mathcal{L}}$ be an equivalent 1D representation of the centroid lattice $\widehat{\mathcal{L}}$ obtained by row lexicographic ordering. To build an integrated framework for detection and tracking, we augment $\overline{\mathcal{L}}$ with an additional dummy state that represents the absence of the target. For convenience, we assign to the absent state the index $(L+r_i+r_s)\,(M+l_i+l_s)+1$. The final 1D *extended lattice* is

$$\widetilde{\mathcal{L}} = \{\,l\colon 1 \le l \le (L+r_i+r_s)(M+l_i+l_s)+1\}\ . \tag{3}$$

The unknown state at frame n is a 1D random variable, z_n, defined on $\widetilde{\mathcal{L}}$. The target motion on the extended lattice $\widetilde{\mathcal{L}}$ is described by a first-order *hidden Markov model* (HMM) specified by the matrix of *transition probabilities*

$$P_T(i,\,j) = P(z_n = i \mid z_{n-1} = j) \tag{4}$$

for all $(i,\,j) \in \widetilde{\mathcal{L}} \times \widetilde{\mathcal{L}}$.

III. Optimal Bayes Detector/Tracker

We make the following assumptions in the derivation of the algorithm:

1. The sequence of clutter frames $\{\mathbf{V}_k\}$, $k \ge 0$, is independent, identically distributed (i.i.d.).
2. The sequence of target states $\{z_k\}$, $k \ge 0$, is statistically independent of the sequence of clutter frames, $\{\mathbf{V}_k\}$, $k \ge 0$.

Let $\mathbf{Y}_0^n = \{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_n\}$, where $\mathbf{y}_k$, $0 \le k \le n$, is a 1D long vector representation of the observed 2D sensor frame $\mathbf{Y}_k$ at instant k. With the assumptions made in the previous paragraph, the posterior probability mass function $P(z_n \mid \mathbf{Y}_0^n)$, $z_n \in \widetilde{\mathcal{L}}$, is computed recursively by the following two-step algorithm (see [8] for further details):

Prediction Step Combining the theorem of total probability and Bayes law, we write

$$P(z_n \mid \mathbf{Y}_0^{n-1}) = \sum_{z_{n-1}} P(z_n \mid z_{n-1},\, \mathbf{Y}_0^{n-1}) \times P(z_{n-1} \mid \mathbf{Y}_0^{n-1})\,. \tag{5}$$

Using the assumption that the current target state, z_n, is statistically independent of the previous clutter frames, $\{\mathbf{V}_k\}$, $0 \le k \le n-1$, and recalling that the sequence of target states $\{z_k\}$, $k \ge 0$, is described by a first-order Markov chain model, it results that, conditioned on z_{n-1}, z_n is statistically independent of the previous observations, $\mathbf{Y}_0^{n-1}$, i.e.,

$$P(z_n \mid z_{n-1},\, \mathbf{Y}_0^{n-1}) = P(z_n \mid z_{n-1})\,. \tag{6}$$

Equation (5) is then rewritten as

$$P(z_n \mid \mathbf{Y}_0^{n-1}) = \sum_{z_{n-1}} P(z_n \mid z_{n-1})\, P(z_{n-1} \mid \mathbf{Y}_0^{n-1})\,. \tag{7}$$

Filtering Step Using Bayes Law,

$$P(z_n \mid \mathbf{Y}_0^n) = C_n p(\mathbf{y}_n \mid z_n, \mathbf{Y}_0^{n-1}) \times P(z_n \mid \mathbf{Y}_0^{n-1}) \tag{8}$$

where C_n is a normalization constant. From the assumption that the sequence of clutter frames $\{\mathbf{V}_k\}$, $k \ge 0$, is i.i.d, we conclude that, conditioned on z_n, the current observation vector $\mathbf{y}_n$ is statistically independent of the past observations, $\mathbf{Y}_0^{n-1}$. Equation (9) reduces then to

$$P(z_n \mid \mathbf{Y}_0^n) = C_n p(\mathbf{y}_n \mid z_n)\, P(z_n \mid \mathbf{Y}_0^{n-1})\,. \tag{9}$$

We now consider detection and tracking.

Detection Let $L_1 = (L+r_i+r_s)\,(M+l_i+l_s)$. Denote by H_0 the hypothesis that the target is absent and, by H_1, the hypothesis that the target is present. Assuming equal cost for misses and false alarms and zero cost for correct decisions, the minimum probability of error detector is the test

$$P(H_0 \mid \mathbf{Y}_0^n) \underset{H_1}{\overset{H_0}{\gtrless}} P(H_1 \mid \mathbf{Y}_0^n)$$

$$\Leftrightarrow \quad \frac{P(z_n = L_1+1 \mid \mathbf{Y}_0^n)}{1-P(z_n = L_1+1 \mid \mathbf{Y}_0^n)} \underset{H_1}{\overset{H_0}{\gtrless}} 1 \,. \tag{10}$$

Tracking Introduce, for all $l \in \overline{\mathcal{L}}$, the conditional probability

$$\begin{aligned} Q_l^f\,[n] &= P(\mathbf{z}_n = l \mid \text{target is present}, \mathbf{Y}_0^n) \\ &= \frac{P(z_n = l \mid \mathbf{Y}_0^n)}{1-P(z_n = L_1+1 \mid \mathbf{Y}_0^n)} \end{aligned} \tag{11}$$

where $\overline{\mathcal{L}}$ is the 1D equivalent centroid lattice, see section II. The maximum a posteriori (MAP) estimate of the the target's centroid position assuming that the target is present is

$$\hat{z}_{\text{map}}\,[n] = \arg\max_{l \in \overline{\mathcal{L}}} Q_l^f\,[n] \,. \tag{12}$$

IV. Approximate ML Parameter Estimation

Let $\mathbf{V}$ be a first order, noncausal $L \times M$ Gauss-Markov random field. We derive next an approximate maximum likelihood algorithm to estimate from $\mathbf{V}$ the corresponding parameters β_h, β_v, and σ_u^2. For the Gauss-Markov model, the negative log-likelihood function is given by [9]

$$\begin{aligned} \mathrm{L}(\mathbf{V}) &= \frac{L\,M}{2}\ln(2\pi) - L\,M \ln\sigma_u \\ &+ \frac{1}{2}\ln\mid \mathbf{A}\mid - \frac{1}{2\sigma_u^2}(\mathbf{V}^T\mathbf{A}\mathbf{V}) \end{aligned} \tag{13}$$

where $\mathbf{A}$ is a block-tridiagonal, block-Toeplitz, $L\,M \times L\,M$ matrix with structure

$$\mathbf{A} = \mathbf{I}_L \otimes \mathbf{I}_M - \beta_h \mathbf{I}_L \otimes \mathbf{H}_M - \beta_v \mathbf{H}_L \otimes \mathbf{I}_M \,. \tag{14}$$

In (14), $\mathbf{I}_r$ is the r x r identity matrix, and $\mathbf{H}_r$ is an r x r matrix whose entries $H_r(i,j)$ are equal to 1 if $\mid i-j \mid = 1$ and equal to zero otherwise. The symbol $\otimes$ denotes the Kronecker or tensor product [10]. From the structure of $\mathbf{A}$, it results that

$$\mathbf{V}^T\mathbf{A}\mathbf{V} = S_v - \beta_h\chi_h - \beta_v\chi_v \tag{15}$$

$$S_v = \sum_{i=1}^{L}\sum_{j=1}^{M} V_{i,j}^2 \tag{16}$$

$$\chi_h = \sum_{i=1}^{L}\sum_{j=1}^{M-1} V_{i,j}\,V_{i,j+1} \tag{17}$$

$$\chi_v = \sum_{i=1}^{L-1}\sum_{j=1}^{M} V_{i,j}\,V_{i+1,j} \,. \tag{18}$$

On the other hand, as shown in [11], the eigenvalues of the matrix $\mathbf{A}$ are

$$\lambda_{ij} = 1 - 2\beta_h\cos(\frac{\pi\,j}{M+1}) - 2\beta_v\cos(\frac{\pi\,i}{L+1}) \tag{19}$$

for $1 \le i \le L, 1 \le j \le M$. The natural logarithm of the determinant of $\mathbf{A}$ on the righthand side of equation (13) can be then computed as

$$\ln\mid \mathbf{A} \mid = \sum_{i=1}^{L}\sum_{j=1}^{M}\ln(\lambda_{ij}) \,. \tag{20}$$

The minimization of the exact negative log-likelihood function in (13) with respect to the parameters β_h, β_v, and σ_u^2 is a computationally intensive nonlinear optimization problem [9]. A suboptimal solution is obtained using the Taylor series approximation

$$\ln(1-\eta) = -\eta - \frac{\eta^2}{2} + o(\eta^2) \,. \tag{21}$$

Combining the approximation in (21) with the trigonometric properties

$$\sum_{i=1}^{N}\cos(\frac{i\,\pi}{N+1}) = 0 \tag{22}$$

$$\sum_{i=1}^{N}\cos^2(\frac{i\,\pi}{N+1}) = \frac{N-1}{2}, \tag{23}$$

we can write

$$\ln\mid A \mid \approx L\,(M-1)\,\beta_h^2 + (L-1)\,M\,\beta_v^2 \,. \tag{24}$$

Substituting equations (15) and (24) into (13), taking the derivatives with respect to β_h and β_v and making them equal to zero, we get

$$\widehat{\beta_h} = \frac{\chi_h}{\sigma_u^2\,L\,(M-1)} \tag{25}$$

$$\widehat{\beta_v} = \frac{\chi_v}{\sigma_u^2\,(L-1)\,M} \,. \tag{26}$$

The Taylor series approximation in (21) is only valid for $\eta << 1$, i.e., for weakly correlated fields. In practice, however, real clutter data tends to be highly correlated, particularly in optical imagery. It is necessary therefore to adjust equations (25) and (26) to accomodate for η not being small. Balram and Moura observed [9] that, for highly correlated GMrfs, the parameters β_h and β_v tend to be located near the boundaries of the parameter space $\mathcal{P}$ that makes $\mathbf{A}$ a positive definite matrix. These boundaries are defined by the equation [9]

$$\mid \beta_h \mid \cos \frac{\pi}{M+1} + \mid \beta_v \mid \cos \frac{\pi}{L+1} = \frac{1}{2} \, . \qquad (27)$$

An heuristic approach [9] to obtain approximate expressions for β_h and β_v in the highly correlated field case is to divide equation (26) by equation (25), thus obtaining a ratio that is independent of σ_u^2, and then to impose the constraint in (27) for $\beta_h = \widehat{\beta_h}$ and $\beta_v = \widehat{\beta_v}$. This procedure leads to the following estimates:

$$\widehat{\beta_h} = \frac{\epsilon \chi_h}{\mid \chi_v \mid \cos(\frac{\pi}{L+1}) + \alpha \mid \chi_h \mid \cos(\frac{\pi}{M+1})}$$

$$\widehat{\beta_v} = \frac{\epsilon \chi_v}{\mid \chi_v \mid \cos(\frac{\pi}{L+1}) + \alpha \mid \chi_h \mid \cos(\frac{\pi}{M+1})}$$

where $\epsilon = 0.5$ and

$$\alpha = \frac{(L-1)M}{L(M-1)} \, .$$

In order to ensure that the estimated parameters $\widehat{\beta_h}$ and $\widehat{\beta_v}$ are valid points in the parameter space $\mathcal{P}$, we empirically replace $\epsilon = 0.5$ with $\tilde{\epsilon} = 0.5 - \delta$, where δ is a small number (for example, $\delta = 10^{-3}$).

The approximate ML estimate of the clutter power, $\widehat{\sigma_u^2}$ can in turn be obtained by taking the derivative of the approximate negative log-likelihood function

$$\tilde{\mathrm{L}}(\mathbf{V}) = L \ln \sigma_u^2 - \frac{1}{2} \left[L(M-1)\beta_h^2 + (L-1)M\beta_v^2 \right] + \frac{1}{2\sigma_u^2} \left[S_x - \beta_h \chi_h - \beta_v \chi_v \right]$$

with respect to σ_u^2, equating it to zero, and making $\beta_h = \widehat{\beta_h}$ and $\beta_v = \widehat{\beta_v}$. Following these steps, the final expression for $\widehat{\sigma_u^2}$ is

$$\widehat{\sigma_u^2} = \frac{1}{L\,M} (S_x - \widehat{\beta}_h \, \chi_h - \widehat{\beta}_v \, \chi_v) \, . \qquad (28)$$

V. Simulation Results

Figure 1(a) shows a 120 × 120 gray-level real intensity image [12] of a snow-covered field in Stockbridge, NY, obtained by an airborne 0.85 μm down-looking active laser radar [5] mounted to a Gulfstream G-1 aircraft. Brighter areas indicate stronger laser returns. We added to the imagery an artificial target template that simulates a military vehicle (tank). The target template is shown isolated in Figure 1(b) as a binary image. The pixel intensity of the target was set so that there is little contrast between the vehicle and the background. To test tracking performance, we simulated a random trajectory for the target template in the background. The target departs from an *unknown* location in the 120 × 120 grid and moves with a constant mean velocity of 2 pixels per frame in both the horizontal and vertical directions. The actual target displacement is a 2D first order random walk fluctuation around the average drift, i.e., if (i, j) is the expected target centroid position according to its deterministic velocity, the real position may be $(i-1, j)$, $(i+1, j)$, $(i, j+1)$, or $(i, j-1)$, each with a uniform 20 % probability. In addition to the simulated target, we also add low power white Gaussian noise to the image sequence.

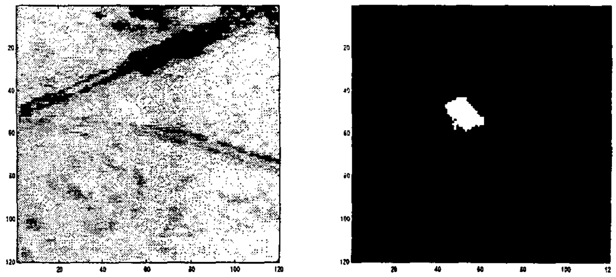
(a) (b)

Fig. 1. Simulated target in cluttered real background: (a) target plus clutter, (b) target template.

We initially preprocess each frame by segmenting it and removing the spatially variant local mean. We then fit a zero-mean first order Gauss-Markov random field model to each frame by estimating the corresponding parameters β_h, β_v, and σ_u^2 using the approximate maximum likelihood estimation algorithm described in section IV. The estimated parameters are subsequently used in the filtering step of the Bayes detector/tracker at each frame. We compare the tracking results for a sequence of 27 frames using (a) the proposed Bayes tracker, and (b) a standard 2D image correlator associated to a linearized Kalman-Bucy filter. The results are shown in Figure 2. The Bayes tracker assumes a uniform initial target position distribution over the entire sensor grid. The linear filter, on the other hand, is initially favored by using a Gaussian initial position prior that is centered in the vicinity of the true initial position and has a small variance. The real simulated trajectory is shown in solid line. The position estimates generated by the Bayes tracker are indicated by the symbol '+', whereas the estimates generated by the linearized Kalman-Bucy filter are interpolated using dashed lines. At each frame, the Bayes algorithm is testing for the presence or absence of target using the detection test (10), before it estimates the target's position using the MAP estimator in (12). In the first half

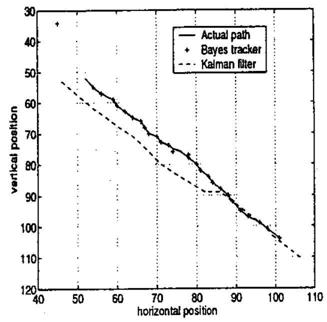

Fig. 2. Nonlinear Bayes detector/tracker versus linearized Kalman-Bucy filter: performance comparison

of the trajectory shown in Figure 2, the simulated tank is going through a heavily cluttered section of the background, and the single frame standard image correlator is unable to track the target. The Kalman-Bucy filter tends to discard the correlator's position estimates and through the inertia in its prediction step, tries to fit a straight line trajectory. In the second half of the simulation, when the tank is on an open field, the image correlator is capable of correctly locating the target and the filtering step of the Kalman-Bucy filter slowly forces the estimated trajectory to approach the true trajectory. By contrast, the Bayes tracker, which has no prior knowledge of the initial position, makes a large initial localization error (the isolated '+' on the top left corner of Figure 2), but, afterwards, as new frames become available, the tracker immediately acquires the target and tracks it almost perfectly. A comparison shows that, even in steady state, the localization error for the Bayes tracker is lower than for the Kalman-Bucy filter, while the acquisition time is much shorter.

VI. Conclusion

We discussed in this paper an optimal Bayesian approach to integrated detection and tracking of randomly moving targets in spatially correlated cluttered image sequences. The algorithm consists of a multiframe minimum probability of error Bayes detector and a multiframe maximum a posteriori (MAP) target position estimator. The detector/tracker design fully incorporates the models for target signature, target motion, and clutter and uses recursive spatio-temporal processing across all available frames to make detection decisions and to generate position estimates. A simulation with an artificial target template added to real clutter background shows the proposed algorithm outperforms the association of a standard single frame image correlator and a linearized Kalman-Bucy filter in a scenario with a dim target.

ACKNOWLEDGMENT

The work of the first author was supported by FAPESP, São Paulo, Brazil. The authors would also like to thank Professor Amir Asif and Dr. Susan Thornton for valuable discussions on clutter parameter estimation.

References

[1] Y. Bar-Shalom and X. Li.*Multitarget-Multisensor Tracking: Principles and Techniques*. YBS, Storrs, CT, 1995.

[2] H. Im and T. Kim, "Optimization of multiframe target detection schemes", *IEEE Transactions on Aerospace and Electronic Systems*, vol.35, n.1, pp 176-187, January 1999.

[3] S. C. Pohlig, "An algorithm for detection of moving optical targets," *IEEE Transactions on Aerospace an Electronic Systems*, vol. 25, pp 56-63, January 1989.

[4] J. Y. Chen and I. S. Reed, "A detection algorithm for optical targets in clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 23, pp 46-59, January 1987.

[5] J. K. Bounds. *The Infrared Airbone Radar Sensor Suite*. RLE Technical Report No. 610, Cambridge, Massachusetts, December 1996.

[6] J. M. F. Moura and N. Balram, "Recursive structure of noncausal Gauss-Markov random fields," *IEEE Transactions on Information Theory*, vol.38, n.2, pp 334-354, March 1992.

[7] M. G. S. Bruno and J. M. F. Moura,"Clutter adaptive multiframe detection/tracking of random signature targets," to be presented at *IEEE International Conference on Accoustics, Speech and Signal Processing (ICASSP)*, Salt Lake City, UT, May 7-11, 2001.

[8] M. G. S. Bruno and J. M. F. Moura, "Multiframe detector/tracker: optimal performance," submitted to *IEEE Transactions on Aerospace and Electronic Systems*, under review.

[9] J. M. F. Moura and N. Balram, "Noncausal Gauss-Markov random fields: Parameter Structure and Estimation," *IEEE Transactions on Information Theory*, vol. 39, n.4, pp 1333-1355, July 1993.

[10] A. Graham, *Kronecker Products and Matrix Calculus with Applications*, Ellis Horwood Limited, 1981.

[11] J. M. F. Moura and M. G. S. Bruno, "DCT/DST and Gauss-Markov fields: conditions for equivalence," *IEEE Transactions on Signal Processing*, vol. 46, n. 9, pp 2571-2574, September 1998.

[12] *Center for Imaging Science*, Johns Hopkins University, URL: http://www.cis.jhu.edu.