Image Authentication Using Distributed Source Coding

Yao-Chung Lin, David Varodayan, *Member, IEEE*, and Bernd Girod, *Fellow, IEEE*

***Abstract*—We present a novel approach using distributed source coding for image authentication. The key idea is to provide a Slepian–Wolf encoded quantized image projection as authentication data. This version can be correctly decoded with the help of an authentic image as side information. Distributed source coding provides the desired robustness against legitimate variations while detecting illegitimate modification. The decoder incorporating expectation maximization algorithms can authenticate images which have undergone contrast, brightness, and affine warping adjustments. Our authentication system also offers tampering localization by using the sum-product algorithm.**

***Index Terms*—Distributed source coding, EM algorithm, image authentication, sum-product algorithm.**

I. Introduction

MEDIA content can be efficiently delivered through intermediaries, such as peer-to-peer (P2P) file sharing and P2P multicast streaming. Popular P2P file sharing systems include BitTorrent, eMule, and KaZaA. In these systems, each user not only receives the requested content but also acts as a relay forwarding the received portions to the other users. Since the same content can be re-encoded several times, media content in those P2P file sharing systems is available in various digital formats, such as JPEG and JPEG2000 for images, and MPEG-1, MPEG-2, and H.264/AVC for videos. On the other hand, the untrusted intermediaries might tamper with the media for a variety of reasons, such as interfering with the distribution of particular files, piggybacking unauthentic content, or generally discrediting a particular distribution system. A 2005 survey indicates that more than 50% of popular songs in KaZaA are corrupted [1], e.g., replaced with noise or different songs. Distinguishing legitimate encoding versions from maliciously tampered ones is important in applications that deliver media content through untrusted intermediaries. The problem is more challenging if some legitimate adjustments, such as cropping and resizing an image, are allowed in addition to lossy compression. Additional adjustments might not change the meaning of the content, but could be misclassified as tampering. Users might also be interested in localizing tampered regions. Distinguishing legitimate encodings with possible adjustments from tampering and localizing tampering are the challenges addressed in this paper. We apply distributed source coding and statistical methods to solve the image authentication problem.

Section II reviews past approaches in image authentication, the fundamentals of distributed source coding, and related work in secure biometrics. Section III introduces the image authentication system using distributed source coding. We formulate image authentication problem as a hypothesis testing problem. The original image projection is quantized and encoded using Slepian–Wolf coding, a form of distributed source coding [2]. By correctly choosing the size of the Slepian–Wolf bitstream, it can be decoded using the legitimate image as side information. Section IV presents an extension of the basic scheme to authenticate images that have undergone legitimate editing, such as contrast, brightness, and affine warping adjustments. The authentication decoder learns the editing parameters directly from the target image through decoding the authentication data using an expectation maximization (EM) algorithm. Section V extends the authentication system to localize tampering in the image.

II. Background

A. Previous Work in Image Authentication

Past approaches for image authentication fall into three groups: forensics, watermarking, and robust hashing. In digital forensics, the user verifies the authenticity of an image solely by checking the received content [3]–[5]. Unfortunately, without any information from the original, one cannot completely confirm the integrity of the received content because content unrelated to the original may pass forensic checking. Another option for image authentication is watermarking. A semi-fragile watermark is embedded into the host signal waveform without perceptual distortion [6]–[8]. Users can confirm authenticity by extracting the watermark from the received content. The system design should ensure that the watermark survives lossy compression, but that it breaks as a result of malicious manipulations. Unfortunately, watermarking authentication is not backward compatible with previously encoded contents; i.e., unmarked content cannot be authenticated later. Embedded watermarks might also increase the bit rate required when compressing a media file.

This paper develops authentication techniques based on robust hashing, which is inspired by cryptographic hashing [9]. In this technique, the user checks the integrity of the received content using a small amount of data derived from the original content. Many hash-based image authentication systems achieve ro-

Manuscript received January 20, 2011; revised April 28, 2011; accepted May 03, 2011. Date of publication May 23, 2011; date of current version December 16, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Chun-Shien Lu.

Y.-C. Lin was with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: yclin79@stanfordalumni.org).

D. Varodayan is with Hewlett-Packard Labs, Palo Alto, CA 94304 USA (e-mail: varodayan@hp.com).

B. Girod is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: bgirod@stanford.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2011.2157515

bustness against lossy compression by using compression-invariant features, such as [10]–[19]. These compression-inspired features are designed for particular compression schemes but fail under other coding schemes or common image processing. Robustness is increased using more sophisticated features, such as block-based histograms [20], zero-mean low-pass Gaussian pseudo-random projection [21], [22], block standard deviations and means [23], [24], column and row projections [25], and transform coefficients [26], [27]. Any fixed projection has the weakness that an attacker who knows the null space of the projection can alter the image without affecting the authentication data. Using pseudo-random projections or tilings, such as in [28], keeps the null space a secret. Similar considerations apply to features calculated in a nonlinear manner. Features robust against rotation, cropping, resizing, or translation have been proposed based on the Radon transform [29]–[31], the Fourier transform [32], and pixel statistics along radii [33]–[35]. Other methods include features important to the human visual system [36]–[42].

Quantization and compression of authentication data has not been studied in depth. Most approaches use coarse quantization. For example, Fridrich *et al.* use 1-bit quantization for random projection coefficients [21], [22], [40], and the relation-based approaches [10]–[12], [14]–[17] can be considered as 1-bit quantizations of coefficient differences. The first to consider error-correcting coding in reducing the image authentication data size were Venkatesan *et al.* [28]. The idea is to project the binary feature vectors of both images into syndrome bits of an error-correcting code and directly compare the syndrome bits to decide the authenticity. The approach of Sun *et al.* uses systematic Hamming codes to obtain the parity check bits of the binary feature vectors as the authentication data [43]. These parity check bits are concatenated with the binary feature vector of the received image to correct the errors introduced by image processing, such as compression. Our novel ideas make further improvements with the knowledge of distributed source coding and statistical methods. Inspired by our approach, Tagliasacchi *et al.* proposed using Wyner–Ziv coding and compressive sensing for image authentication by exploiting additional assumptions on the sparsity of tampering [44].

B. Lossless Distributed Source Coding

The problem of compressing features X of the original image relative to features Y of the target image is a distributed source coding problem as shown in Fig. 1. Source X is available at the encoder, but the side information Y is available at the decoder only. Slepian and Wolf proved that X can be compressed to a rate $R \geq H(X \mid Y)$ and still be decoded without loss in the presence of Y [2]. Conversely, when R is less than $H(X \mid Y)$, the probability of decoding error will be bounded away from zero.

State-of-the-art practical Slepian–Wolf coding often employs low-density parity-check (LDPC) codes [45], [46]. The work reported in this paper likewise uses LDPC codes and employs them to efficiently encode random projections of images.

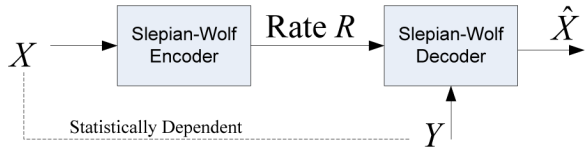


Fig. 1. The source X and side information Y are statistically dependent, but Y is available only at the decoder.

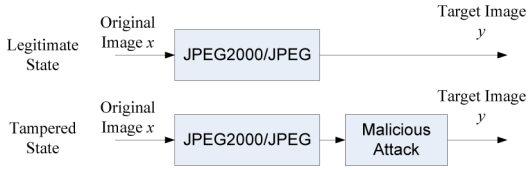


Fig. 2. The target image y is modeled as an output of a two-state lossy channel. In the legitimate state, the channel consists of lossy compression and reconstruction, such as JPEG and JPEG2000; in the tampered state, the channel further applies a malicious attack.

C. Secure Biometrics

Our approach has similarities to Slepian–Wolf coding for secure storage of biometric data reported in [47], [48]. The problem is to robustly hash enrollment versions of the biometric. The idea is to encode features of the enrollment biometric, so that decoding is possible only with a correlated authentication biometric acting as side information. The secure biometric problem and the image authentication problem have important differences. For secure biometrics, the biometric data from two different people are assumed to be independent. In image authentication, the tampered target images are usually correlated to the original but with statistics different to those of the authentic target images. Thus, the secure biometric problem requires hypothesis testing against *independence* under rate constraints [49], while image authentication is a more general rate-constrained hypothesis testing problem [50], [51]. The observation that the target images are usually correlated supports our use of the EM algorithm for learning unknown editing parameters and the sum-product algorithm for tampering localization.

III. Image Authentication System

We can conveniently formulate image authentication as a hypothesis testing problem. The authentication data provides information about the original image to the user. The user makes the authentication decision based on the target image and the authentication data. We first describe a two-state channel that models the target image and then present the image authentication system using distributed source coding.

A. Two-State Channel

We model the target image y using a two-state channel, shown in Fig. 2. In the legitimate state, the channel performs lossy compression and reconstruction, such as JPEG or JPEG2000, with peak signal-to-noise ratio (PSNR) of 30 dB or better. In the tampered state, it includes a malicious attack.

(a) (b) (c)

Fig. 3. Examples of the two-state lossy channel output. (a) x original, (b) y at the output of the legitimate channel, and (c) y at the output of the tampered channel.

Fig. 3 demonstrates a sample input and two outputs of this channel. The source image x is a Kodak test image at 512×512 resolution. In the legitimate state, the channel is JPEG2000 compression and reconstruction at (the worst permissible) 30 dB PSNR. In the tampered state, a further malicious attack is applied: a 19×163 pixel text banner is overlaid on the reconstructed image and some objects are removed.

The joint statistics of x and y vary depending on the state of the channel. In the legitimate state, the difference resembles white noise due to the compression; in the tampered state, the channel additionally introduces tampering which results in image-like differences in some regions. This suggests that low frequency components can greatly distinguish legitimate and tampered regions. Let X and Y be low-frequency block projections of images x and y, respectively. The image authentication problem at the projection level in the hypothesis testing setting is described as follows:

$$X \,|\, Y \sim \begin{cases} P(X \,|\, Y) = \mathcal{N}(Y, \sigma_0^2) \\ Q(X \,|\, Y) = (1-\gamma)\mathcal{N}(Y, \sigma_0^2) + \gamma P_{\text{tamp}}(X \,|\, Y) \end{cases} \tag{1}$$

where the distribution is $P(X \,|\, Y)$ if y is legitimate and $Q(X \,|\, Y)$ if it is tampered. Also, $\gamma \in [0, 1]$ is the fraction of tampered image blocks, and $P_{\text{tamp}}(X \,|\, Y)$ is their probability model. We assume that $P_{\text{tamp}}(X \,|\, Y) = U(X)$ is a uniform distribution over the dynamic range of X. Having both projections X and Y, the optimal decision is based on the likelihood ratio test: $P(X,Y)/Q(X,Y) \gtrless T$. The next section describes our image authentication scheme which uses these statistical assumptions to generate authentication data using distributed source coding.

B. Proposed Image Authentication System

In our authentication system shown in Fig. 4, a pseudorandom projection (based on a randomly drawn seed K_s) is applied to the original image x and the projection coefficients X are quantized to yield X_q. The authentication data are comprised of two parts, both derived from X_q. The Slepian–Wolf bitstream $S(X_q)$ is the output of a Slepian–Wolf encoder based on LDPC codes [45] and the much smaller digital signature $D(X_q, K_s)$ consists of the seed K_s and a cryptographic hash value of X_q signed with a private key.

The authentication data are generated by a server upon request. Each response uses a different random seed K_s, which is provided to the decoder as part of the authentication data. This prevents an attack which simply confines the tampering to the nullspace of the projection. Based on the random seed, for each 16×16 nonoverlapping block B_i, we generate a 16×16 pseudorandom matrix P_i by drawing its elements independently from a Gaussian distribution $\mathcal{N}(1, \sigma_p^2)$ and normalizing so that $\|P_i\|_2 = 1$. We choose $\sigma_p = 0.2$ empirically. In this way, we maintain the properties of the mean projection while gaining sensitivity to high-frequency attacks. The inner product $\langle B_i, P_i \rangle$ is uniformly quantized into an element of X_q.

The rate of the Slepian–Wolf bitstream $S(X_q)$ determines how statistically similar the target image must be to the original to be declared authentic. If the conditional entropy $H(X_q \,|\, Y)$ exceeds the bitrate R in bits per pixel, X_q cannot be decoded correctly [2]. Therefore, the rate of $S(X_q)$ should be chosen to be just sufficient to authenticate the legitimate image at its worst permissible quality. In our system, we select a Slepian–Wolf bitrate just sufficient to authenticate both legitimate 30 dB JPEG2000 and JPEG reconstructed versions of the original image. Practically, the Slepian–Wolf bitrate is determined by finding the minimum decodable rate for the training images with the worst permissible quality. This worst permissible quality is an external parameter that depends on the particular application. Generally, if a smaller quality degradation is permissible, fewer bits are required for authentication. If a worse quality is permissible, more bits are needed.

At the receiver, the user seeks to authenticate the image y with authentication data $S(X_q)$ and $D(X_q, K_s)$. It first projects y to Y in the same way as during authentication data generation using the same random seed K_s. A Slepian–Wolf decoder reconstructs $X_q{}'$ from the Slepian–Wolf bitstream $S(X_q)$ using Y as side information. Decoding is via LDPC belief propagation [45] initialized according to the statistics of the legitimate channel state at the worst permissible quality for the given original image. Finally, the image digest of $X_q{}'$ is computed and compared to the image digest, decrypted from the digital signature $D(X_q, K_s)$ using a public key. If these two image digests do not match, the receiver recognizes that image y is tampered. Otherwise the receiver makes a decision based on the likelihood ratio test: $P(X_q{}', Y)/Q(X_q{}', Y) \lessgtr T$, where P and Q are probability models derived from (1) for legitimate and tampered states, respectively, and T is a fixed decision threshold.

The authentication system presented in this section can address various types of lossy compression. The next section discusses an adaptive distributed source coding decoder to broaden the robustness of the system for some common adjustments, such as contrast and brightness adjustment, and affine warping.

IV. Learning Unknown Parameters of Image Adjustment

It is not uncommon that a target image has undergone additional adjustments besides compression. Some of these we might want to accept as legitimate image adjustments. For example, the image might be slightly cropped and resized to meet the size and resolution of the client display or contrast and brightness adjustment may have been adjusted for an image that is too dark or to bright. If we consider those image adjustment legitimate, the basic image authentication system described in the previous section would fail; even a slight resizing or brightness or contrast change would be considered tampering.

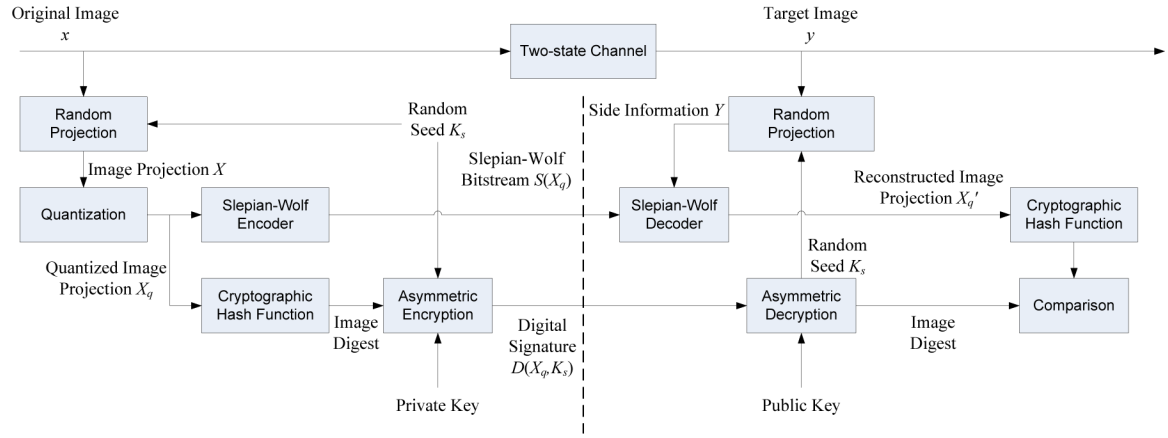


Fig. 4. Image authentication system using distributed source coding. The authentication data consists of a Slepian–Wolf encoded quantized pseudorandom projection of the original image, a random seed, and a signature of the image projection. The target image is modeled as an output of the two-state lossy channel shown in Fig. 2. The user projects the target image using the same projection to yield the side information and tries to decode the Slepian–Wolf bitstream using the side information. If the decoding fails, i.e., the hash value of the reconstructed image projection does not match the signature, the verification decoder claims it is tampered, otherwise, the reconstructed image projection along with the side information is examined using hypothesis testing.

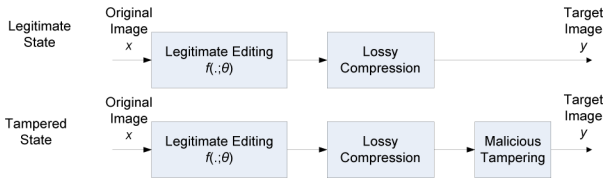


Fig. 5. The target image is modeled as an output of a two-state channel affected by a global editing function $f(.;\theta)$ with unknown but fixed parameter θ. In the tampered state, the channel additionally applies malicious tampering.

Fig. 6. One of the Kodak test images. (a) The original image and (b) a legitimate image with contrast increased by 20%, brightness decreased by 10/255, and rotated 5 degrees around the center. The target image (b) is compressed and reconstructed by JPEG at 30 dB PSNR. (c) Realigned target image color overlaid. The blue areas associated with the 16×16 blocks indicate the cropped-out regions; the other blocks form the cropped-in region.

Decoding the authentication data by trying out all possible editing parameters is clearly not feasible, the computational complexity would be overwhelming.

In the following, we present a novel solution in which the authentication decoder learns the editing parameters directly from the target image through decoding the authentication data using an expectation maximization (EM) algorithm. We introduce a two-state channel with unknown editing parameters to formulate the problem and an EM decoder for images that have simultaneously undergone contrast, brightness, and affine warping adjustment.

A. Two-State Channel With Unknown Adjustment Parameters

We model the target image by way of a two-state channel with unknown adjustment parameters as shown in Fig. 5. In both states, the channel adjusts the image via legitimate editing with a fixed but unknown parameter θ. In the legitimate state, we model $y = f(x;\theta) + z$, where x and y are the original and the target images, respectively, and z is noise introduced by compression and reconstruction. In the tampered state, the channel additionally applies malicious tampering.

Fig. 6 demonstrates the channel for a Kodak test image at 512×512 resolution. Fig. 6(b) shows a target image which has simultaneously undergone contrast, brightness, and affine warping adjustment: $y(\mathbf{m}) = f(x;\mathbf{A},\mathbf{b},\alpha,\beta) = \alpha x(\mathbf{n}) + \beta$, where $\mathbf{n} = \mathbf{A}\mathbf{m} + \mathbf{b}$, $\mathbf{n},\mathbf{m} \in R^2$ are the corresponding coordinates in the original and target images, respectively, $\alpha,\beta \in R$ are contrast and brightness adjustment parameters, and $\mathbf{A} \in R^{2\times 2}, \mathbf{b} \in R^2$ are transformation and translation parameters, respectively. In this case, there are 8 scalar parameters. Exhaustive search is not practical. Moreover, since the authenticity decision is based on likelihood ratio test: $P(X_q,y;\theta)/Q(X_q,y;\theta) \gtrless T$, accurate estimation of θ is needed for confident decision results.

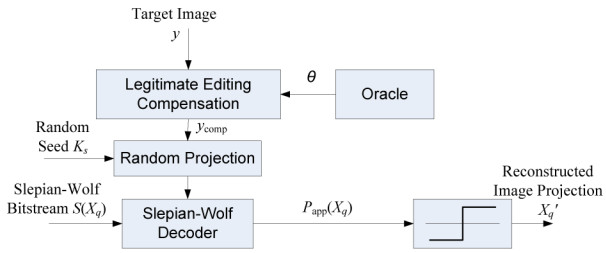


Fig. 7. The oracle decoder knows the parameters and compensates the target image to align with the authentication data. Then the Slepian–Wolf is decoded using the compensated target image as side information to yield an *a posteriori* pmf of the quantized projection $P_{\text{app}}(X_q)$. The reconstructed quantized image projection is the result of a hard decision on $P_{\text{app}}(X_q)$.

Fig. 7 shows a decoder that has access to an oracle knowing the true editing parameters of the target image. The target image is compensated using the parameters provided by the oracle, and the decoder decodes the Slepian–Wolf bitstream

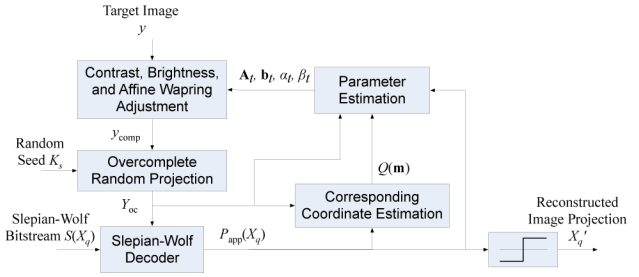


Fig. 8. The Slepian–Wolf decoder with contrast, brightness, and affine warping adjustment learning decodes the Slepian–Wolf bitstream $S(X_q)$ using the target image y. Each iteration produces soft estimation of corresponding coordinates $\mathbf{m}$ and quantized original projections X_q in the E-step and updates the adjustment parameters in the M-step.

and tests the target image and reconstructed image projection in the same way described in Section III. The authentication decision is based on the reconstructed image projection and the compensated target image. Due to affine warping and cropping, some portions of the original image are cropped out in the target image y. The cropped-out areas of the target image are not considered in the authentication decision. Fig. 6(c) shows the target image realigned to the original. The blue areas in Fig. 6(c) indicate the cropped-out regions. We refer to the remaining area of the image as the "cropped-in" region. Clearly, the oracle decoder is not practical, but it will be useful as an upper performance bound later on. Next we show how to turn the oracle decoder into a practical one using statistical learning techniques.

B. EM Decoder for Contrast, Brightness, and Affine Warping Adjustment

We consider a target image that has simultaneously undergone contrast, brightness and affine warping adjustment. The contrast of the example target image shown in Fig. 6(b) is increased by 20%, and brightness decreased by 10/255. It is then rotated counterclockwise by 5 degrees around the image center, cropped to 512×512 and JPEG compressed and reconstructed at 30 dB PSNR. Recall that we model the editing as $y(\mathbf{m}) = \alpha x(\mathbf{A}\mathbf{m} + \mathbf{b}) + \beta + z(\mathbf{m})$, where

$$\mathbf{A} = \begin{bmatrix} 0.996 & -0.087 \\ 0.087 & 0.996 \end{bmatrix}$$

and

$$\mathbf{b} = \begin{bmatrix} 23 \\ -21 \end{bmatrix}$$

for a 5-degree counterclockwise rotation and cropping, and $\alpha = 1.2$, $\beta = -10$, for contrast and brightness changes.

Unlike past approaches in which the projection or the features might be invariant to the contrast, brightness, and affine warping adjustment, we solve this problem by decoding the authentication data while learning the parameters that establish the correlation between the target and original images. Estimation of the adjustment parameters requires the target image and the original image projections, but the latter is not available before decoding. This situation with latent variables to estimate can be addressed using EM.

The EM Slepian–Wolf decoder in Fig. 8 decodes the Slepian–Wolf bitstream $S(X_q)$ using the target image y and yields the reconstructed image projection $X_q{}'$. The E-step updates the *a posteriori* probability mass function (pmf) $P_{\text{app}}(X_q)$ and estimates corresponding coordinates for a subset of reliably-decoded projections. The M-step updates the affine warping parameters based on the corresponding coordinate distributions, denoted $Q(\mathbf{m})$ in Fig. 8. This loop of EM iterations terminates when hard decisions on $P_{\text{app}}(X_q)$ satisfy the constraints imposed by $S(X_q)$.

In the iteration t, the E-step fixes the parameters $\mathbf{A}_t$, $\mathbf{b}_t$, α_t, and β_t at their current hard estimates and obtains a compensated image y_{comp}. We derive intrinsic pmfs for the image projections X_q as follows. In the cropped-in region, we use Gaussian distributions centered at the random projection values of y_{comp}, and in the cropped-out region, we use uniform distributions. Then we run three iterations of LDPC decoding on the *a priori* pmfs with the Slepian–Wolf bitstream $S(X_q)$ to produce *a posteriori* pmfs $P_{\text{app}}(X_q(i) = x_q)$.

We estimate the corresponding coordinates $\mathbf{m}^{(i)}$ for those projections for which $\max_{x_q} P_{\text{app}}(X_q(i) = x_q) > T = 0.995$, denoting this set of reliably-decoded projection indices as $\mathcal{C}$.[1] We also denote the maximizing reconstruction value x_q to be $x_q^{\max}(i)$. For the projection $X_q(i)$, we produce the pmf $P_{\text{app}}(\mathbf{m}^{(i)} = \mathbf{m}; \mathbf{A}_t, \mathbf{b}_t, \alpha_t, \beta_t)$ by matching $X_q(i)$ to the overcomplete projections Y_{oc} of y_{comp} through $\mathbf{m}$ over a small search window. Specifically, $P_{\text{app}}(\mathbf{m}^{(i)} = \mathbf{m}; \mathbf{A}_t, \mathbf{b}_t, \alpha_t, \beta_t)$ is proportional to the integral over the quantization interval of $x_q^{\max}(i)$ of a Gaussian centered at the projection of a block at $\mathbf{m}$ in the image y_{comp}. Since y_{comp} in the later iterations is closer to the original image, we empirically set the search window size to $[2W(t)+1]^2$, where $W(t) = \max(40 \times 0.92^t, 1)$, and the variance for the Gaussian to $\max(100 \times 0.92^{2t}, 4)$. The update of the latent variable $\mathbf{m}$ is written as

$$Q_i(\mathbf{m}) := P\left(\mathbf{m}^{(i)} = \mathbf{m} \,|\, x_q^{\max}(i), y, \mathbf{n}^{(i)}; \mathbf{A}_t, \mathbf{b}_t, \alpha_t, \beta_t\right).$$

In the M-step, we estimate the parameters $\mathbf{A}'$, $\mathbf{b}'$, α', and β' with respect to Y_{oc} by holding the corresponding coordinate pmfs $Q_i(\mathbf{m})$ fixed and maximizing a lower bound of the log-likelihood function:

$$\begin{aligned}
&L(\mathbf{A}', \mathbf{b}', \alpha', \beta') \\
&\quad \equiv \sum_{i \in \mathcal{C}} \log P\left(x_q^{\max}(i), \mathbf{n}^{(i)}, Y_{\text{oc}}; \mathbf{A}', \mathbf{b}', \alpha', \beta'\right) \\
&\quad = \sum_{i \in \mathcal{C}} \log \sum_{\mathbf{m}^{(i)}} P\left(x_q^{\max}(i), \mathbf{n}^{(i)}, Y_{\text{oc}}, \mathbf{m}^{(i)}; \mathbf{A}', \mathbf{b}', \alpha', \beta'\right) \\
&\quad \geq \sum_{i \in \mathcal{C}} \sum_{\mathbf{m}} Q_i(\mathbf{m}) \log P\left(x_q^{\max}(i), \mathbf{n}^{(i)}, Y_{\text{oc}} \,|\, \mathbf{m}; \mathbf{A}', \mathbf{b}', \alpha', \beta'\right) \\
&\quad = \sum_{i \in \mathcal{C}} \sum_{\mathbf{m}} Q_i(\mathbf{m}) \log P(\mathbf{n}^{(i)} \,|\, \mathbf{m}; \mathbf{A}', \mathbf{b}') \\
&\qquad + \sum_{i \in \mathcal{C}} \sum_{\mathbf{m}} Q_i(\mathbf{m}) \log P\left(x_q^{\max}(i), Y_{\text{oc}} \,|\, \mathbf{m}, \mathbf{n}^{(i)}; \alpha', \beta'\right).
\end{aligned}$$

The lower bound is due to Jensen's inequality and concavity of $\log(\cdot)$. Note also that $P(x_q^{\max}(i), Y_{\text{oc}} \,|\, \mathbf{m}, \mathbf{n}^{(i)}; \alpha', \beta')$ does not

[1]To guarantee that $\mathcal{C}$ is nonempty, we make sure to encode a small portion of the quantized image projection X_q with degree-1 syndrome bits. The decoder knows those values with probability 1 and includes their indices in $\mathcal{C}$.

depend on the parameters $\mathbf{A}'$ and $\mathbf{b}'$, and $P(\mathbf{n}^{(i)} \,|\, \mathbf{m}; \mathbf{A}', \mathbf{b}')$ does not depend on the parameters α' and β'. Thus, we can maximize the lower bound separately over these two sets of parameters. The affine warping parameters are updated using (2) derived from the least squares method with assumption that $P(\mathbf{n}^{(i)} \,|\, \mathbf{m}; \mathbf{A}', \mathbf{b}')$ is a Gaussian with mean at $\mathbf{A}'\mathbf{m} + \mathbf{b}'$.

$$\begin{bmatrix} A'_{11} & A'_{21} \\ A'_{12} & A'_{22} \\ b'_1 & b'_2 \end{bmatrix} := E[G^T G]^{-1} E[G^T] \begin{bmatrix} \vdots & \vdots \\ n_1^{(i)} & n_2^{(i)} \\ \vdots & \vdots \end{bmatrix} \tag{2}$$

where

$$G = \begin{bmatrix} \cdots & m_1^{(i)} & \cdots \\ \cdots & m_2^{(i)} & \cdots \\ \cdots & 1 & \cdots \end{bmatrix}^T$$

and

$$E[G^T G] = \sum_{i \in \mathcal{C}} \begin{bmatrix} E\left[\left(m_1^{(i)}\right)^2\right] & E\left[m_1^{(i)} m_2^{(i)}\right] & E\left[m_1^{(i)}\right] \\ E\left[m_1^{(i)} m_2^{(i)}\right] & E\left[\left(m_2^{(i)}\right)^2\right] & E\left[m_2^{(i)}\right] \\ E\left[m_1^{(i)}\right] & E\left[m_2^{(i)}\right] & 1 \end{bmatrix}.$$

Similarly, we model $P(X_q^{\max}(i) \,|\, Y_{\text{oc}}, \mathbf{m}, \mathbf{n}^{(i)}; \alpha', \beta')$ as a quantized Gaussian with mean at $(Y_{\text{oc}}(\mathbf{m}) - \beta')/(\alpha')$. Setting partial derivatives with respect to α' and β' to zero, we obtain the updates:

$$\alpha' := \frac{|\mathcal{C}| \sum_{i \in \mathcal{C}} \mu^i_{XY} - \sum_{i \in \mathcal{C}} \mu^i_X \sum_{j \in \mathcal{C}} \mu^j_Y}{|\mathcal{C}| \sum_{i \in \mathcal{C}} \mu^i_{X^2} - \left(\sum_{i \in \mathcal{C}} \mu^i_X\right)^2}$$

$$\beta' := \frac{1}{|\mathcal{C}|} \sum_{i \in \mathcal{C}} \mu^i_Y - \alpha' \mu^i_X$$

where

$$\begin{aligned} \mu^i_X &= E_{\mathbf{m} \sim Q_i}\left[E\left[X | Y_{\text{oc}}(\mathbf{m}), x_q^{\max}(i)\right]\right], \\ \mu^i_Y &= E_{\mathbf{m} \sim Q_i}[Y_{\text{oc}}(\mathbf{m})], \\ \mu^i_{X^2} &= E_{\mathbf{m} \sim Q_i}\left[E[X^2 \,|\, Y_{\text{oc}}(\mathbf{m}), x_q^{\max}(i)\right], \\ \mu^i_{XY} &= E_{\mathbf{m} \sim Q_i}\left[Y_{\text{oc}}(\mathbf{m}) E\left[X \,|\, Y_{\text{oc}}(\mathbf{m}), x_q^{\max}(i)\right]\right]. \end{aligned}$$

Note that the parameters, $\mathbf{A}', \mathbf{b}', \alpha', \beta'$, are with respect to y_{comp}. The parameters with respect to the target image y for the next iteration are updated as follows: $\mathbf{A}_{t+1} = \mathbf{A}_t \mathbf{A}'$, $\mathbf{b}_{t+1} = \mathbf{A}_t \mathbf{b}' + \mathbf{b}_t$, $\alpha_{t+1} = \alpha_t \alpha'$, and $\beta_{t+1} = \alpha_t \beta' + \beta_t$.

The likelihood ratio test for authenticity is $P(X_q, y; \mathbf{A}, \mathbf{b}, \alpha, \beta)/Q(X_q, y; \mathbf{A}, \mathbf{b}, \alpha, \beta) \gtrless T$, measured over the cropped-in area of the compensated target image where $\mathbf{A}, \mathbf{b}, \alpha, \beta$ are the final estimated parameters with respect to y.

Fig. 9 demonstrates the efficiency of the EM decoder by illustrating the traces of parameter searching for different decoders facing contrast and brightness changes. The ground truth of the contrast parameter is 0.84, and brightness is 10. The oracle decoder directly outputs the ground truth. The decoder unaware of adjustment uses 1 and 0 for contrast and brightness parameters, respectively. In Fig. 9(c), the exhaustive search decoder tries to decode the authentication data using samples in the parameter space from -0.75 to 1.2 of contrast parameter and -20 to 20 of brightness parameter until it obtains a parameter sample that can successfully decode the bitstream. The discrete search space makes the resulting parameters inaccurate and the computational complexity grows exponentially as the parameter dimension increases. Fig. 9(d) shows the search trace of our proposed EM decoder. Even though the initial parameters are far from the ground truth, the decoder approaches it in a manageable number of iterations. Unlike exhaustive search, the EM decoder estimates the parameters in a continuous space.

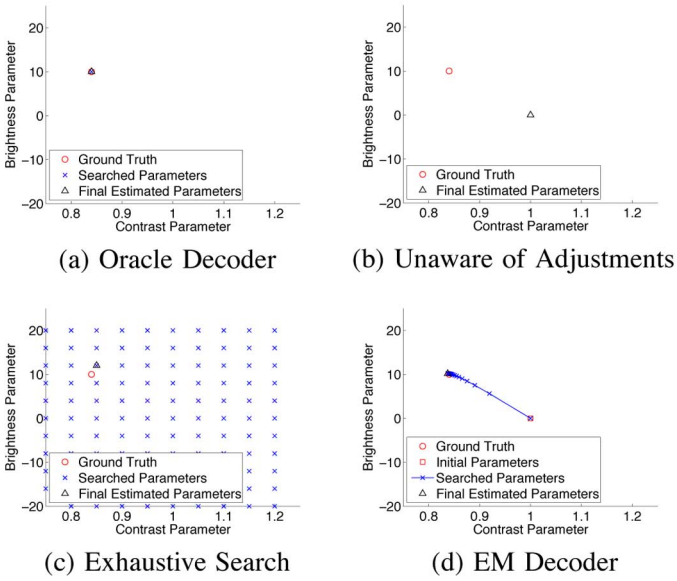


Fig. 9. Search traces for different decoders. (a) The oracle decoder directly outputs the ground truth; (b) the decoder unaware of adjustment outputs (1,0) for contrast and brightness parameters; (c) the exhaustive search decoder tries to decode the authentication data using the parameters in the discrete search space, until it reaches a parameter that can successfully decode the authentication data; (d) the proposed EM decoder iteratively updates the parameters and decodes the authentication data.

The proposed EM decoder can handle slight manipulations including slight downsampling and cropping. If the manipulation is too severe (such as 90 degree rotation), the system will deem the target image as tampered. Possible ways to handle severe manipulations include normalizing the original and target images [52] or starting with a set of images obtained from the target image (e.g., all of its 90 degree rotations).

The decoding complexity is $O(MW)$, where M is the number of projection coefficients, and W is the search window size. In the E-step, computing $Q(\mathbf{m})$ takes $O(W)$ per projection coefficient. In the M-step, the computation of moments for parameter estimation also takes $O(W)$ per projection coefficient.

Our system decodes the authentication data using legitimate target images that may have undergone contrast, brightness, and affine warping adjustments. The next section considers decoding with tampered target images as side information.

V. Tampering Localization

Localization of tampering requires reconstructing the original image projection using the tampered image as side information. As will be shown in simulation results, using legitimate

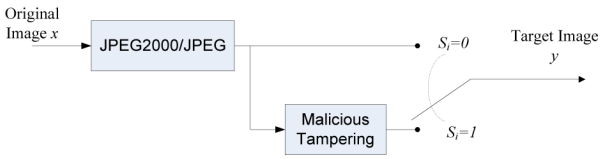
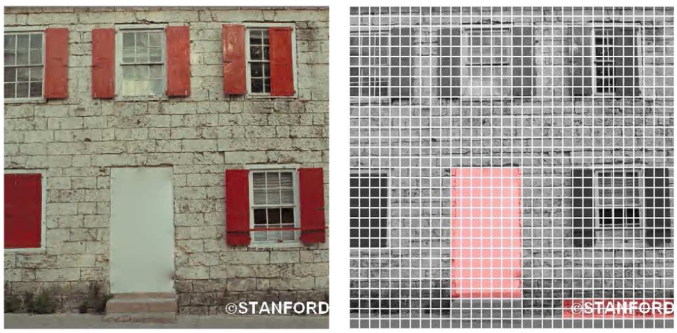


Fig. 10. Space-varying two-state lossy channel. The image is divided into nonoverlapping blocks. Each block has an associated channel state indicating whether the block is tampered or legitimate.



(a) (b)

Fig. 11. The target image in (a) is a tampered version of the original image in Fig. 3(a). The image in (b) is the overlaid channel state for each 16×16 block. The red blocks are tampered, and the others are legitimate.

editing models to decode the authentication data with tampered side information needs a high authentication data rate. In this section, we describe a localization decoder that requires a much lower authentication data rate. The decoder handles the correlation between the original image and slightly tampered target images using a sum-product algorithm over a factor graph [53]. We first formulate the localization problem using a space-varying two-state channel and then describe the localization decoder factor graph.

A. Space-Varying Two-State Channel

The space-varying two-state channel is shown in Fig. 10. In the legitimate state, the channel output is legitimate editing, such as JPEG2000 compression and reconstruction. The tampered state additionally includes malicious tampering. The channel state variable S_i is defined per nonoverlapping 16×16 block of image y. If any pixel in block B_i is part of the tampering, $S_i = 1$; otherwise, $S_i = 0$. The authentication problem discussed in Sections III and IV is a decision per image; the tampering localization problem can be formulated as deciding on S_i for each block, given the Slepian–Wolf bitstream $S(X_q)$. Fig. 11(b) shows the channel states overlaid on a tampered target image shown in Fig. 11(a). The red blocks are tampered, and the others are legitimate.

Given the quantized original image projection X_q, and the target image projection Y, one can infer the channel state S using Bayes' theorem:

$$P(S \mid X_q, Y) = \frac{P(X_q, Y \mid S)P(S)}{P(X_q, Y)}. \tag{3}$$

The localization decoder requires more information than the authentication decoder since it additionally estimates the channel

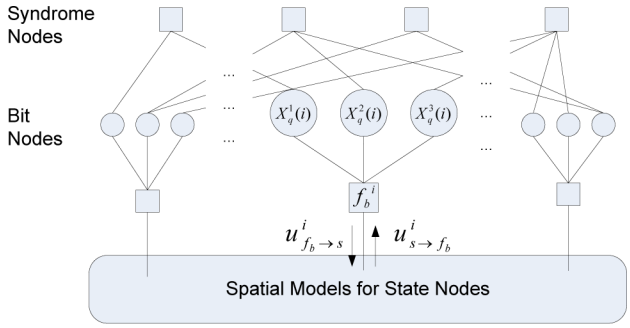


Fig. 12. Factor graph for the localization decoder.

states, and a tampered image is usually less correlated to X_q than an authentic one. If authentication is run before tampering localization, the localization decoder can reuse the authentication data and merely request incremental localization data. Such an implementation is possible using rate-adaptive LDPC codes [46]. In practice, the bitrate of the incremental localization data is estimated using a representative training set of tampered images. Next we introduce the decoder factor graph that connects the LDPC decoding to the channel state inference. The sum-product algorithm over the factor graph simultaneously decodes the Slepian–Wolf bitstream and localizes the tampering.

B. Decoder Factor Graph

A factor graph [53] is a bipartite graphical model that represents a factorization of a joint probability distribution of random variables. There are two classes of nodes: the variable nodes represent the random variables of interest; the factor nodes represent the probabilistic relationships among the adjacent variable nodes. Based on the factor graph representation, the sum-product algorithm efficiently marginalizes the approximate joint distribution for all variables.

The factor graph in Fig. 12 shows the relationship among the Slepian–Wolf bitstream (at syndrome nodes), the image projection X_q (quantized to 3 bits at bit nodes), and the side information and channel states (within the spatial model). The variable nodes of interest are $[X_q^1(i), X_q^2(i), X_q^3(i)]$ which form the binary representation of $X_q(i)$ and the channel states S_i contained in the spatial model. The factor node at each syndrome node is an indicator function of the satisfaction of that syndrome constraint. The factor $f_b^i(X_q(i), S_i) = P(X_q(i)|Y(i); S_i)$ represents the relationship between image projection $X_q(i)$, side information Y_i, and the channel state S_i. When $S_i = 0$, factor $f_b^i(X_q(i), 0)$ is proportional to the integral of a Gaussian distribution with mean $Y(i)$ and a fixed variance σ_z^2 over the quantization interval of $X_q(i)$. When $S_i = 1$, $f_b^i(X_q(i), 1)$ is uniform. The spatial model of the channel states is independent and identically distributed (IID), a 1D Markov chain, or a 2D Markov random field. Decoding is via the sum-product algorithm executed over the entire factor graph. The decision about the value of state S_i is a threshold operating on the resulting marginal probability. Details of the algorithm are presented in [54], [55].

VI. Simulation Results

We use test images at 512×512 resolution in 8-bit gray scale resolution. The authentic test images are JPEG or JPEG2000

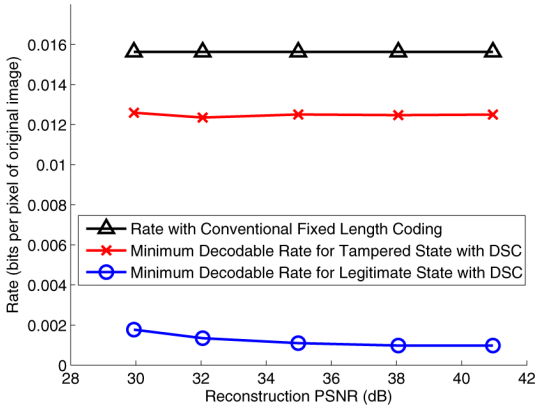


Fig. 13. Minimum rates (averaged for the tampered states) for correctly decoding Slepian–Wolf bitstream for the image *Lena* with the projection X quantized to 4 bits.

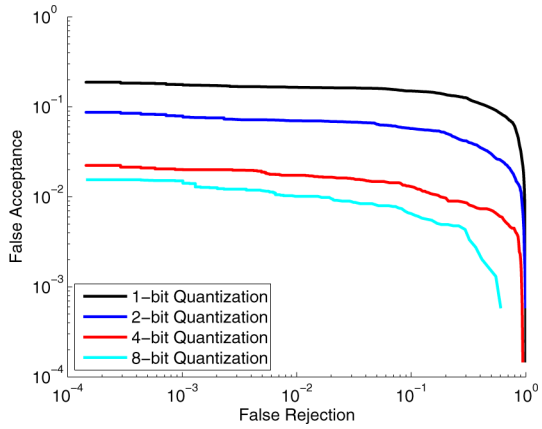


Fig. 14. Receiver operating characteristic curves of tampering detection with different number of bits in quantization of X for test images. This demonstrates that higher quantization precision offers better detection performance.

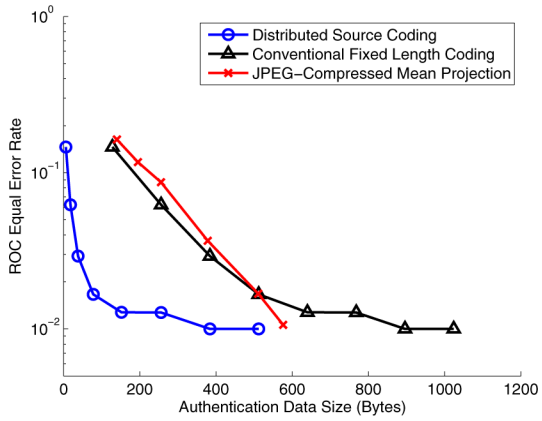


Fig. 15. ROC equal error rates for different authentication data sizes using conventional fixed length coding, distributed source coding, and JPEG-compressed mean projection.

compressed and reconstructed at several qualities. The malicious attack consists of the overlay of text banners at a random location in the image or removing a randomly selected Maximally Stable Extremal Region (MSER) [56] of 1500 pixels of larger by interpolating the region. For the text banners, the text color is white or black, whichever is more visible, to avoid generating trivial attacks, such as white text on a white area.

Using this data set, we demonstrate the performance of the authentication system for compressed images, the authentication system with EM decoder for adjusted images, and the tampering localization system.

A. Authentication of Compressed Images

The quantization of the authentication encoder is varied so that the Slepian–Wolf encoder processes between 1 to 8 bits, starting with the most significant. The Slepian–Wolf codec is implemented using rate-adaptive LDPC codes [46] with block size of 1024 bits. During authentication data generation, the bitplanes of X are encoded successively. The bitplanes are conditionally decoded, with each decoded bitplane acting as additional side information for subsequent bitplanes [57].

Fig. 13 compares the minimum decodable rates of the Slepian–Wolf bitstream $S(X_q)$ for *Lena* with the projection X quantized to 4 bits. The following observations also hold for other images and levels of quantization. The rate required to decode $S(X_q)$ with legitimately created side information is significantly lower than the rate (averaged over 100 trials) when the side information is tampered, for JPEG2000 or JPEG reconstruction PSNR above 30 dB. Moreover, as the PSNR increases, the rate for legitimate side information decreases, while the rate for tampered side information stays high and close to the conventional fixed length coding. The rate gap justifies our choice for the Slepian–Wolf bitstream size: the size just sufficient to authenticate both legitimate 30 dB JPEG2000 and JPEG reconstructed versions of the original image.

We now fix the authentication data sizes of different numbers of bits in quantization to evaluate the tampering detection using 3 450 legitimate and 3 450 tampered test images with $\sigma_0 = 2$ and $\gamma = 0.0233$ in (1) for legitimate and tampered models. We measure the false acceptance rate (the chance that a tampered image is falsely accepted as a legitimate one) and the false rejection rate (the chance that a legitimate image is falsely detected as a tampered one). Fig. 14 compares the receiver operating characteristic (ROC) curves for tampering detection with different numbers of bits in quantization by sweeping the decision threshold T in the likelihood ratio test.

Fig. 14 shows that higher quantization precision offers better detection performance, but at the cost of more authentication data. Fig. 15 plots the ROC equal error rate versus the authentication data size and demonstrates that distributed source coding reduces the data size by more than 80% compared to conventional fixed length coding at an equal error rate of 2%. Distributed source coding also outperforms a baseline authentication based on JPEG. The encoder of this system uses JPEG to compress the coefficients of a 16×16-block mean projection. The decoder's decision is based on $\max_i |X'(i) - Y(i)|$, where X' is the reconstructed original image projection and Y is the image projection of the target image.

B. Authentication of Adjusted Images

Now we evaluate the performance of the EM decoder for the test images with affine warping adjustments. The first experiment shows the minimum decodable rates for rotated and

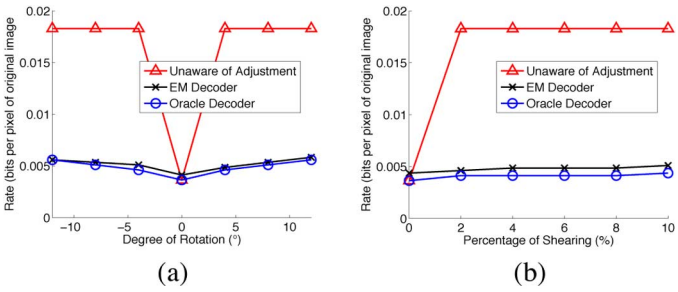


Fig. 16. Minimum rate for decoding authentication data using legitimate adjusted test images as side information for different using different decoders. (a) The test images have undergone rotation. (b) The test images have undergone horizontal shearing. The EM decoder requires minimum rates only slightly higher than the oracle decoder, while the decoder unaware of adjustment requires higher and higher rate as the adjustment increases.

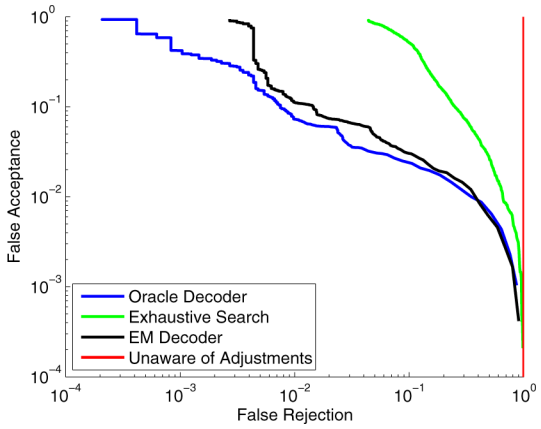


Fig. 17. Receiver operating characteristic curves for different decoders. The target images have undergone random contrast, brightness, and affine warping adjustments and JPEG/JPEG2000 compression. The EM decoder performance is very close to that of the oracle decoder, while the decoder unaware of adjustments rejects authentic test images with high probability. The exhaustive search decoder, which tries parameter samples at intervals of 1 for $\mathbf{b}$, 0.1 for β, and 0.01 for the others rounded from the ground truth, also suffers from high probability of false rejection due to the inaccurate parameters used.

sheared target images. We apply an affine warping adjustment to the images and crop them to 512×512. Then JPEG2000 or JPEG compression and reconstruction are applied at 30 dB reconstruction PSNR. In the tampered state, the malicious attack overlays a 20×122 pixel text banner randomly on the image. The image projection X is quantized to 4 bits, and the Slepian–Wolf encoder uses a 4096-bit LDPC code with 400 degree-1 syndrome nodes. Fig. 16 compares the minimum rates for decoding $S(X_q)$ with legitimate test images using three different decoding schemes: the EM decoder that learns the affine parameters, an oracle decoder that knows the parameters, and a decoder unaware of adjustment that always assumes no adjustment. Fig. 16(a) and (b) show the results when the affine warping adjustments are rotation around the image center and horizontal shearing, respectively. The EM decoder requires minimum rates only slightly higher than the oracle decoder, while the decoder unaware of adjustment requires higher and higher rates as the adjustment increases.

For the next experiment, we set the authentication data size to 250 bytes and measure false acceptance and rejection rates. The acceptance decision is made based on the likelihood of X_q and y with estimated parameters within the estimated cropped-in blocks. The settings remain the same except that parameter α is randomly drawn from $[0.9, 1.1]$, β from $[-10, 10]$, A_{11} and A_{22} from $[0.95, 1.05]$, A_{21} and A_{12} from $[-0.05, 0.05]$, and b_1 and b_2 from $[-10, 10]$. The JPEG2000/JPEG reconstruction PSNR is selected from 30 to 42 dB. With 15,000 trials, Fig. 17 shows the receiver operating characteristic curves. The EM decoder performance is very close to that of the oracle decoder, while the decoder unaware of adjustments rejects authentic test images with high probability. The exhaustive search decoder, which tries parameter samples at intervals of 0.01 for $\mathbf{A}$ and α, 0.1 for β, and 1 for $\mathbf{b}$ rounded from the ground truth, also suffers from high probability of false rejection due to the inaccurate parameters used. In the legitimate case, the EM decoder estimates the transform parameters A_{11}, A_{21}, A_{12}, A_{22}, b_1, b_2, α, and β with mean squared error 4.5×10^{-7}, 2.6×10^{-6}, 3.4×10^{-7}, 1.6×10^{-6}, 0.05, 0.54, 2.0×10^{-5}, and 0.34, respectively.

C. Tampering Localization

In practice, the localization decoder would only run if the authentication decoder deems an image to be tampered, so we test the tampering localization system only with maliciously tampered images. We use test images with JPEG2000 or JPEG compression and reconstruction applied at several qualities above 30 dB. The malicious tampering consists of the overlaying of up to five text banners of different sizes at random locations in the image. The text banner sizes are 198×29, 29×254, 119×16, 16×131, and 127×121 pixels. The text color is white or black, depending on which is more visible, again avoiding generating trivial attacks, such as overlaying white text on a white area. All five text banners are placed for malicious tampering, because greater tampering makes tampering more easily detected, but makes localization more difficult.

Fig. 18 shows the Slepian–Wolf bitstream $S(X_q)$ of these rates (in bits per pixel of the original image x) for *Lena* with X_q in 4-bit quantization. The placement of text banners is random for 100 trials, leading to tampering of 12% to 17% of the nonoverlapping 16×16 blocks of the original image x. Decoding the localization data using a legitimate model for tampered target images requires a bit rate close to fixed length coding. Using the localization decoder instead results in 65% less bit rate when the spatial model is IID, and even less rate when the spatial model is 1D or 2D. Fig. 19 shows the ROC curves of undetected tampered pixels against falsely deemed tampered blocks for these spatial models, and demonstrates that the advantage of 1D and 2D spatial models over the IID model is in reducing the rate of undetected tampered pixels.

VII. Conclusions

This paper presents and investigates a novel image authentication scheme that distinguishes legitimate encoding variations of an image from tampered versions based on distributed source coding and statistical methods. A two-state lossy channel model represents the statistical dependency between the original and the target images. Tampering degradations are captured by using a statistical image model, and legitimate compression noise is assumed to be additive white Gaussian noise.

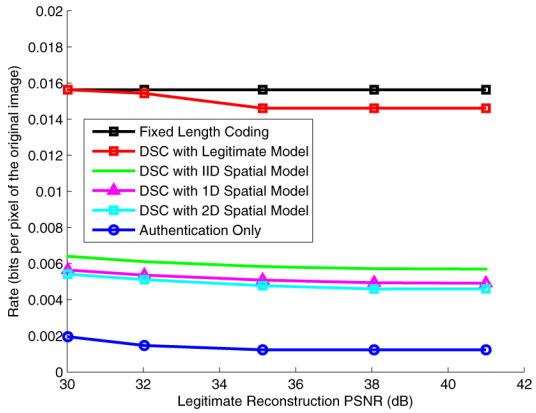


Fig. 18. Minimum rates for decoding Slepian–Wolf bitstream under various spatial models.

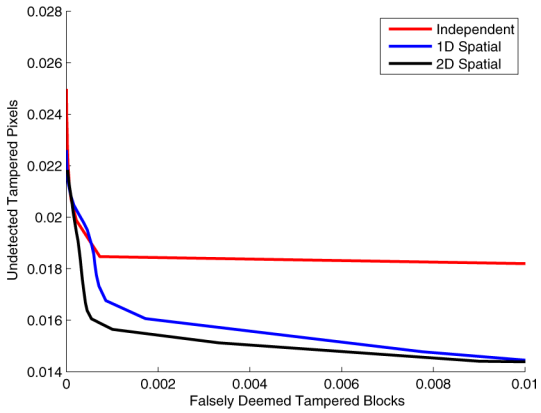


Fig. 19. Receiver operating characteristic curves of the tampering localization decoders using spatial models. The rates of falsely deemed tampered blocks can reach zero, while keeping the undetected tampered pixel rates at about 2%, since most of the blocks falsely deemed untampered have only a few pixels tampered. In most cases, 1D and 2D spatial models achieve a lower undetected tampered pixel rate at a given falsely deemed tampered block rate.

Slepian–Wolf coding that exploits the correlation between the original and the target image projections achieves significant rate savings. The Slepian–Wolf decoder is extended using expectation maximization algorithms to address target images that have undergone contrast, brightness, and affine warping adjustment. The localization decoder infers the tampered locations and decodes the Slepian–Wolf bitstream by applying the sum-product algorithm over a factor graph which represents the relationship among the Slepian–Wolf bitstream, projections of the original image and the target image, and the block states. Spatial models are applied to exploit the spatial correlation of the tampering. Distributed source coding is an ideal tool for the image authentication problem in which the data sent for authentication are highly correlated to the information available at the receiver.

REFERENCES

[1] J. Liang, R. Kumar, Y. Xi, and K. W. Ross, "Pollution in P2P file sharing systems," in *Proc. IEEE Infocom*, Mar. 2005, vol. 2, pp. 1174–1185.

[2] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.

[3] H. Farid, "Image forgery detection," *IEEE Signal Process. Mag.*, vol. 26, no. 2, pp. 16–25, Mar. 2009.

[4] J. Lukas and J. Fridrich, "Estimation of primary quantization matrix in double compressed JPEG images," presented at the Digital Forensic Research Workshop, Cleveland, OH, Aug. 2003.

[5] A. Popescu and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3948–3959, Oct. 2005.

[6] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for images, audio and video," in *Proc. IEEE Int. Conf. Image Process.*, Lausanne, Switzerland, Sep. 1996.

[7] J. J. Eggers and B. Girod, "Blind watermarking applied to image authentication," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process.*, Salt Lake City, UT, May 2001.

[8] R. B. Wolfgang and E. J. Delp, "A watermark for digital images," in *Proc. IEEE Int. Conf. Image Process.*, Lausanne, Switzerland, Sep. 1996.

[9] W. Diffie and M. E. Hellman, "New directions in cryptography," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 6, pp. 644–654, Jan. 1976.

[10] C.-Y. Lin and S.-F. Chang, "Generating robust digital signature for image/video authentication," in *ACM Multimedia: Multimedia and Security Workshop*, Bristol, U.K., Sep. 1998, pp. 49–54.

[11] C.-Y. Lin and S.-F. Chang, "A robust image authentication method surviving JPEG lossy compression," in *Proc. SPIE Conf. Storage and Retrieval for Image and Video Database*, San Jose, CA, Jan. 1998.

[12] C.-Y. Lin and S.-F. Chang, "A robust image authentication method distinguishing JPEG compression from malicious manipulation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 2, pp. 153–168, Feb. 2001.

[13] C. Kailasanathan, R. S. Naini, and P. Ogunbona, "Compression tolerant DCT based image hash," in *Proc. Int. Conf. Distributed Computing Syst. Workshops*, May 2003, pp. 562–567.

[14] C.-S. Lu and H.-Y. M. Liao, "Structural digital signature for image authentication: An incidental distortion resistant scheme," in *ACM Workshops on Multimedia*, Los Angeles, CA, 2000, pp. 115–118.

[15] C.-S. Lu and H.-Y. M. Liao, "Structural digital signature for image authentication: An incidental distortion resistant scheme," *IEEE Trans. Multimedia*, vol. 5, no. 2, pp. 161–173, Jun. 2003.

[16] M. Abdel-Mottaleb, G. Vaithilingam, and S. Krishnamachari, "Signature-based image identification," in *Proc. SPIE Conf. Multimedia Syst. Applicat.*, Boston, MA, Sep. 1999, pp. 22–28.

[17] J. Oostveen, T. Kalker, and J. Haitsma, "Visual hashing of video: Applications and techniques," in *Proc. SPIE Conf. Applicat. of Digital Image Process.*, San Diego, CA, Jul. 2001, pp. 121–131.

[18] F. Ahmed and M. Siyal, "A secure and robust hashing scheme for image authentication," in *Proc. Int. Conf. Information, Communications, and Signal Process.*, 2005, pp. 705–709.

[19] M. Schlauweg, D. Pröfrock, and E. Müller, "JPEG2000-based secure image authentication," in *Workshop on Multimedia and Security*, Geneva, Switzerland, 2006, pp. 62–67.

[20] M. Schneider and S.-F. Chang, "A robust content based digital signature for image authentication," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 1996, vol. 3, pp. 227–230.

[21] J. Fridrich, "Robust bit extraction from images," in *Int. Conf. Multimedia Computing and Syst.*, Jul. 1999, vol. 2, pp. 536–540.

[22] J. Fridrich and M. Goljan, "Robust hash functions for digital watermarking," in *Int. Conf. Inf. Technol.: Coding and Computing*, 2000, pp. 178–183.

[23] C. Kailasanathan and R. C. Naini, "Image authentication surviving acceptable modifications using statistical measures and k-mean segmentation," in *Workshop on Nonlinear Signal and Image Process.*, Jun. 2001.

[24] D.-C. Lou and J.-L. Liu, "Fault resilient and compression tolerant digital signature for image authentication," *IEEE Trans. Consumer Electronics*, vol. 46, no. 1, pp. 31–39, Feb. 2000.

[25] L. Xie, G. R. Arce, and R. F. Graveman, "Approximate image message authentication codes," *IEEE Trans. Multimedia*, vol. 3, no. 2, pp. 242–252, Jun. 2001.

[26] R.-X. Zhan, K. Y. Chau, Z.-M. Lu, B.-B. Liu, and W. H. Ip, "Robust image hashing for image authentication based on DCT-DWT composite domain," in *Proc. IEEE Int. Conf. Intelligent Syst. Design and Applicat.*, Nov. 2008, vol. 2, pp. 119–122.

[27] H. Zhang, H. Zhang, Q. Li, and X. Niu, "Predigest Watson's visual model as perceptual hashing method," in *Int. Conf. Convergence and Hybrid Inf. Technol.*, Nov. 2008, vol. 2, pp. 617–620.

[28] R. Venkatesan, S.-M. Koon, M. H. Jakubowski, and P. Moulin, "Robust image hashing," in *Proc. IEEE Int. Conf. Image Process.*, 2000, vol. 3, pp. 664–666.

[29] F. Lefebvre, J. Czyz, and B. Macq, "A robust soft hash algorithm for digital image signature," in *Int. Conf. Multimedia and Expo*, Baltimore, MD, 2003.
[30] J. S. Seo, J. Haitsma, T. Kalker, and C. D. Yoo, "Affine transformation resilient image fingerprinting," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process.*, Hong Kong, China, 2003.
[31] H.-L. Zhang, C.-Q. Xiong, and G.-Z. Geng, "Content based image hashing robust to geometric transformations," in *Proc. Int. Symp. Electronic Commerce and Security*, May 2009, vol. 2, pp. 105–108.
[32] A. Swaminathan, Y. Mao, and M. Wu, "Robust and secure image hashing," *IEEE Trans. Inf. Forensics and Security*, vol. 1, no. 2, pp. 215–230, Jun. 2006.
[33] C. De Roover, C. De Vleeschouwer, F. Lefebvre, and B. Macq, "Robust video hashing based on radial projections of key frames," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 4020–4037, Oct. 2005.
[34] S. Yang, "Robust image hash based on cyclic coding the distributed features," in *Proc. Int. Conf. Hybrid Intelligent Syst.*, Aug. 2009, vol. 2, pp. 441–444.
[35] Z. Tang, S. Wang, X. Zhang, and W. Wei, "Perceptual similarity metric resilient to rotation for application in robust image hashing," in *Proc. Int. Conf. Multimedia and Ubiquitous Eng.*, Jun. 2009, pp. 183–188.
[36] V. Monga and B. L. Evans, "Perceptual image hashing via feature points: Performance evaluation and tradeoffs," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3452–3465, Nov. 2006.
[37] S. Bhattacharjee and M. Kutter, "Compression tolerant image authentication," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 1998, vol. 1, pp. 435–439.
[38] J. Dittmann, A. Steinmetz, and R. Steinmetz, "Content-based digital signature for motion pictures authentication and content-fragile watermarking," in *Proc. IEEE Int. Conf. Multimedia Computing and Syst.*, Jul. 1999, vol. 2, pp. 209–213.
[39] M. P. Queluz, "Towards robust content based techniques for image authentication," in *IEEE Workshop on Multimedia Signal Process.*, Dec. 1998, pp. 297–302.
[40] S. Roy and Q. Sun, "Robust hash for detecting and localizing image tampering," in *Proc. IEEE Int. Conf. Image Process.*, San Antonio, TX, 2007.
[41] C.-S. Lu, C.-Y. Hsu, S.-W. Sun, and P.-C. Chang, "Robust mesh-based hashing for copy detection and tracing of images," in *Proc. IEEE Int. Conf. Multimedia and Expo*, Jun. 2004, vol. 1, pp. 731–734.
[42] M. Schlauweg and E. Müller, "Gaussian scale-space features for semi-fragile image authentication," in *Proc. Picture Coding Symp.*, May 2009, pp. 1–4.
[43] Q. Sun, S.-F. Chang, M. Kurato, and M. Suto, "A new semi-fragile image authentication framework combining ECC and PKI infrastructure," in *Proc. IEEE Int. Symp. Circuits and Syst.*, Phoenix, AZ, May 2002.
[44] M. Tagliasacchi, G. Valenzise, and S. Tubaro, "Hash-based identification of sparse image tampering," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2491–2504, Nov. 2009.
[45] A. Liveris, Z. Xiong, and C. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 440–442, Oct. 2002.
[46] D. Varodayan, A. Aaron, and B. Girod, "Rate-adaptive codes for distributed source coding," *EURASIP Signal Process. J., Special Section on Distributed Source Coding*, vol. 86, no. 11, pp. 3123–3130, Nov. 2006.
[47] E. Martinian, S. Yekhanin, and J. S. Yedidia, "Secure biometrics via syndromes," in *Allerton Conf. Communications, Control and Computing*, Monticello, IL, Sep. 2005.
[48] A. Vetro, S. C. Draper, S. Rane, and J. S. Yedida, *Securing Biometric Data*. Boston, MA: Academic, 2009, ch. 11, pp. 293–324.
[49] R. Ahlswede and I. Csiszar, "Hypothesis testing with communication constraints," *IEEE Trans. Inf. Theory*, vol. 32, no. 4, pp. 533–542, Jul. 1986.
[50] T. S. Han, "Hypothesis testing with multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 33, no. 6, pp. 759–772, Nov. 1987.
[51] T. S. Han and S. Amari, "Statistical inference under multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2300–2324, Oct. 1998.
[52] N. Khanna, A. Roca, G. T. C. Chiu, J. P. Allebach, and E. J. Delp, "Improvements on image authentication and recovery using distributed source coding," in *Proc. SPIE Conf. Media Forensics and Security*, 2009.
[53] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 10, pp. 498–519, Feb. 2001.
[54] Y.-C. Lin, D. Varodayan, and B. Girod, "Spatial models for localization of image tampering using distributed source codes," in *Proc. Picture Coding Symp.*, Lisbon, Portugal, Nov. 2007.
[55] Y.-C. Lin, "Image Authentication Using Distributed Source Coding," Ph.D. dissertation, Stanford University, Stanford, CA, 2010.
[56] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," in *British Machine Vision Conf.*, 2002.
[57] A. Aaron, S. Rane, E. Setton, and B. Girod, "Transform-domain Wyner-Ziv codec for video," in *SPIE Visual Communications and Image Process. Conf.*, San Jose, CA, 2004.

Yao-Chung Lin received the B.S. degree in computer science and information engineering and the M.S. degree in electrical engineering from National Chiao Tung University, Taiwan. He received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 2010.

His research interests include distributed source coding applications, multimedia systems, and video processing and compression.

David Varodayan (M'11) received the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 2005 and 2010, respectively.

He is currently a NSF Corporate Research Postdoctoral Fellow at Hewlett-Packard Laboratories in Palo Alto, CA. His research interests include distributed source coding, image and video processing, and signal processing for the smart grid.

Dr. Varodayan received the EURASIP Signal Processing Journals Most Cited Paper Award in 2009 and Best Student Paper Award on two occasions: IEEE Workshop on Multimedia Signal Processing in 2006 and European Signal Processing Conference in 2007.

Bernd Girod (F'98) received an Engineering Doctorate from the University of Hannover, Germany, and an M.S. degree from the Georgia Institute of Technology, Atlanta, GA.

He is Professor of electrical engineering and (by courtesy) computer science in the Information Systems Laboratory of Stanford University, Stanford, CA, since 1999. Previously, he was a Professor in the Electrical Engineering Department of the University of Erlangen-Nuremberg, Germany. His current research interests are in the areas of video compression, networked media systems, and image-based retrieval. He has published over 450 conference and journal papers, as well as five books. As an entrepreneur, he has been involved with several startup ventures, among them Polycom, Vivo Software, 8x8, and RealNetworks.

Prof. Girod received the EURASIP Signal Processing Best Paper Award in 2002, the IEEE Multimedia Communication Best Paper Award in 2007, the EURASIP Image Communication Best Paper Award in 2008, and the EURASIP Technical Achievement Award in 2004. He is a Fellow of the IEEE, a EURASIP Fellow, and a member of the German National Academy of Sciences (Leopoldina).