Traitor-Within-Traitor Behavior Forensics: Strategy and Risk Minimization

H. Vicky Zhao, *Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

***Abstract*—Multimedia security systems have many users with different objectives and they influence each other's performance and decisions. Behavior forensics analyzes how users with conflicting interests interact with and respond to each other. Such investigation enables a thorough understanding of multimedia security systems and helps the digital rights enforcer offer stronger protection of multimedia. This paper analyzes the dynamics among attackers during multiuser collusion. The colluders share not only the profit from the redistribution of multimedia but also the risk of being detected by the content owner, and an important issue in collusion is fairness of the attack (i.e., whether all attackers share the same risk) (e.g., whether they have the same probability of being detected). While they might agree so, some selfish colluders may break their fair-play agreement in order to further lower their risk. This paper investigates the problem of "traitors within traitors" in multimedia forensics, in an effort to formulate the dynamics among attackers and understand their behavior to minimize their own risk and protect their own interests. As the first work on the analysis of this colluder dynamics, this paper explores some possible strategies that a selfish colluder can use to minimize his or her probability of being caught. We show that processing his or her fingerprinted copy before multiuser collusion helps a selfish colluder further lower his or her risk, especially when the colluded copy has high resolution and good quality. This paper also investigates the optimal precollusion processing strategies for selfish colluders to minimize their risk under the quality constraints.**

***Index Terms*—Behavior forensic, fairness, multiuser collusion, risk minimization, traitors within traitors.**

I. Introduction

SHARING and distributing digital multimedia over networks is becoming popular these days due to recent developments in network and multimedia technologies, and this raises the fundamental and critical issue of protecting multimedia content from illegal alteration and unauthorized redistribution. To trace traitors and identify the source of the illicit copy, the emerging digital fingerprinting technology uniquely labels each distributed copy with identification information. However, the uniqueness of each distributed copy also enables several attackers to collectively mount attacks and remove traces of the identifying fingerprints by combining information from differently fingerprinted copies of the same content [1], [2]. To support multimedia forensics, digital fingerprinting should resist multiuser collusion as well as attacks by a single adversary [3]–[5]. In the literature, techniques from a wide range of disciplines (e.g., error correcting codes [6], finite projective geometry [7], and combinatorial theories [8]) were used to design anticollusion multimedia fingerprints. In group-oriented fingerprint design [9], prior knowledge of the potential collusion pattern was utilized to improve collusion resistance. These prior works explored the unique features of multimedia, jointly considered fingerprint design and embedding, and seamlessly embedded fingerprints into the host signal using the traditional data hiding technique for multimedia.

In multimedia fingerprinting systems, different users have different goals and objectives, and they influence each other's decisions and performance. Behavior forensics formulates the dynamics among attackers during collusion and the dynamics between the colluders and the detector, and investigates how users interact with and respond to each other. Such investigation enables the digital rights enforcer to have a better understanding of the multimedia fingerprinting systems (e.g., how attackers behave during collusion, which information of the collusion can help improve the detection performance, etc.). It helps the digital rights enforcer offer stronger protection of multimedia content.

During multiuser collusion, attackers share not only the profit from the illegal redistribution of multimedia but also the risk of being caught by the digital rights enforcer. Since no colluder is willing to take a larger risk than the others, attackers usually agree to distribute the risk evenly among themselves. Such attacks are referred to as fair collusion attacks. During collusion, each attacker ensures that he or she is not taking a higher risk than the others, and achieving fairness of the attack is an important issue during collusion.

Most prior work in the literature assumed that colluders receive fingerprinted copies of the same quality and emphasized the analysis of collusion strategies and effectiveness. In [10] and [11], collusion attacks were modeled as averaging attacks followed by the addition of noise. In [12], collusion attacks were generalized to linear shift-invariant filtering followed by additive noise. Several types of collusion attacks were studied in [2], including a few nonlinear collusion attacks, and detailed analysis of linear and nonlinear collusion attacks on orthogonal fingerprints was provided in [13]. The Gradient attack was proposed in [14], which uses the combination of several basic nonlinear collusion attacks in [13] during collusion. The work in [15] evaluated the collusion resistance of multimedia fingerprints as a function of system parameters, including fingerprint length, the total number of users, and the system requirements.

Manuscript received June 15, 2005; revised July 28, 2006.

H. V. Zhao is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: vzhao@ece.ualberta.ca).

K. J. R. Liu is with the Department of Electrical and Computer Engineering, Institutes for Systems Research, University of Maryland, College Park, MD 20742 USA (e-mail: kjrliu@eng.umd.edu).

Color versions of Figs. 2, 3, 6, 7, and 9–16 are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2006.885023

The work in [16] investigated how colluders can achieve fairness of collusion when they receive copies of different resolutions due to network and device heterogeneity, and analyzed the constraints on, and the effectiveness of, fair collusion on scalable multimedia fingerprinting.

These prior works on collusion attacks assumed that all colluders keep the agreement to share the same risk during collusion—they provide one another with correct information on their fingerprinted signals and use the copies received from the content owner during collusion. However, the assumption of fair play may not always hold. Some selfish colluders may break the fairness agreement with others and try to further lower their risk of being caught. For example, they may process their fingerprinted signals before collusion and use the processed copies instead of the originally received ones during collusion. When a colluder breaks the fairness agreement, he or she does not necessarily increase others' risk: he or she merely reduces his or her own risk and, therefore, reduces his or her relative risk with respect to the other colluders. In some scenarios, the selfish colluder does increase the other attackers' absolute risk and we call his or her behavior malicious. The existence of selfish colluders raises complex dynamics for multiuser collusion. From the traitor-tracing perspective, it is important to study this problem of traitors within traitors in digital fingerprinting and understand the attackers' behavior during collusion to minimize their own risk and protect their own interest. This investigation of the traitor-within-traitor dynamics helps build a complete model of multiuser collusion and enables to offer stronger protection of multimedia.

As the first work on the analysis of the colluder dynamics, this paper investigates the possible strategies that the selfish colluders can use to minimize their own risk, and evaluates their performance. The rest of the paper is organized as follows. We begin in Section II with the introduction of digital fingerprinting systems and the dynamics among attackers during collusion. In Section III, we investigate the strategies for the selfish colluders to reduce the energy of the embedded fingerprints before multiuser collusion, which further reduce the selfish colluders' probability of being detected. Section IV studies the problem of traitors within traitors in scalable fingerprinting systems when attackers receive copies of different resolutions, and investigates how a selfish colluder can change the resolution of his or her fingerprinted copy to reduce the risk of being captured. Conclusions are drawn in Section V.

II. System Model

A. Digital Fingerprinting Systems for Multimedia Forensics

1) Fingerprint Embedding: Spread-spectrum embedding has been widely used in multimedia fingerprinting systems due to its robustness against many attacks [1], [17]. In additive spread-spectrum embedding for video applications, for the jth frame in the video sequence represented by a vector $\mathbf{S}_j$ of length N_j, the content owner generates a unique fingerprint $\mathbf{W}_j^{(i)}$ of length N_j for each user $\mathbf{u}^{(i)}$ in the system. The fingerprinted copy that is distributed to $\mathbf{u}^{(i)}$ is $X_j^{(i)}(k) = S_j(k) + \mathrm{JND}_j(k) \cdot W_j^{(i)}(k)$, where $X_j^{(i)}(k), S_j(k)$ and $W_j^{(i)}(k)$ are the kth components of the fingerprinted frame $\mathbf{X}_j^{(i)}$, the host signal $\mathbf{S}_j$ and the fingerprint vector $\mathbf{W}_j^{(i)}$, respectively. JND_j is the just-noticeable difference from human visual models [17], and it is used to control the energy and achieve the imperceptibility of the embedded fingerprints. Finally, the content owner transmits to each user $\mathbf{u}^{(i)}$ the fingerprinted frames $\{\mathbf{X}_j^{(i)}\}$.

In this paper, we consider orthogonal fingerprint modulation [15], [8] and assume that the total number of users is much smaller than the length of the embedded fingerprints. With orthogonal modulation, fingerprints for different users are orthogonal to each other and have equal energy (i.e., for user $\mathbf{u}^{(i_1)}$ and $\mathbf{u}^{(i_2)}$)

$$\left\langle \mathbf{W}^{(i_1)}, \mathbf{W}^{(i_2)} \right\rangle = \|\mathbf{w}\|^2 \delta_{i_1,i_2} \tag{1}$$

where δ_{i_1,i_2} is the Dirac–Delta function. It equals to 1 if and only if $i_1 = i_2$ and 0 otherwise. $\|\mathbf{w}_j\|^2$ depends on the fingerprint's length N_j and $\|\mathbf{w}_j\|^2 = N_j \cdot \xi^2$ where ξ is a constant. To resist intracontent collusion attacks on video watermarking [18], [19], in each fingerprinted copy $\{\mathbf{X}_j^{(i)}\}$, the fingerprints $\mathbf{W}_{j_1}^{(i)}$ and $\mathbf{W}_{j_2}^{(i)}$ that are embedded in adjacent frames $\mathbf{S}_{j_1}$ and $\mathbf{S}_{j_2}$, respectively, correlate with each other. The correlation between $\mathbf{W}_{j_1}^{(i)}$ and $\mathbf{W}_{j_2}^{(i)}$ depends on the similarity between the two host frames $\mathbf{S}_{j_1}$ and $\mathbf{S}_{j_2}$, similar to the work in [20].

2) Multiuser Collusion Attacks: During collusion, the colluders collect all of the fingerprinted copies that they received, apply the multiuser collusion function to these copies, and generate a new copy in which the originally embedded fingerprints are removed or attenuated. A recent investigation in [15] showed that under the constraints that the colluded copies from different collusion have the same perceptual quality, the performance of nonlinear collusion attacks is similar to that of the averaging attack. Thus, it suffices to consider averaging-based collusion only.

3) Fingerprint Detection and Colluder Identification: In this paper, we consider a nonblind detection scenario, where the host signal is available to the detector and is first removed from the test copy before fingerprint detection and colluder identification. Once the content owner discovers the existence of an illegal copy in the market, for each frame $\mathbf{V}_j$ in the colluded copy, the detector first extracts the fingerprint $\mathbf{Y}_j = (\mathbf{V}_j - \mathbf{S}_j)/\mathrm{JND}_j$. Then, he or she calculates the similarity between the extracted fingerprint $\{\mathbf{Y}_j\}$ and each original fingerprint $\{\mathbf{W}_j^{(i)}\}$, compares a predetermined threshold h, and outputs the estimated identities of the colluders $\hat{\mathrm{SC}}$.

The correlation-based detector is widely used in the literature to measure the similarity between the extracted fingerprint and the original fingerprint [1], [8], [15]. For each user $\mathbf{u}^{(i)}$, following the thresholding detection strategy in [15], the detector calculates the detection statistics:

$$T_N^{(i)} = \sum_j \left\langle \mathbf{Y}_j, \mathbf{W}_j^{(i)} \right\rangle \Big/ \sqrt{\sum_l \left\| \mathbf{W}_l^{(i)} \right\|^2} \tag{2}$$

where $\|\mathbf{W}_j^{(i)}\|$ is the Euclidean norm of $\mathbf{W}_j^{(i)}$. For a given threshold h, the estimated colluder set is $\hat{\mathrm{SC}} = \{i : T_N^{(i)} > h\}$.

B. Traitors Within Traitors and the Dynamics Among Attackers During Collusion

During collusion, the attackers not only share the profit from the illegal redistribution of multimedia content, they also share the risk of being detected by the digital rights enforcer. Since no one is willing to take a higher risk than the others, they usually apply fair collusion and distribute the risk evenly among themselves. Therefore, fairness is a very important issue that the attackers need to address during collusion. Each attacker makes sure that he or she has the same probability of being detected as others. To achieve fairness of collusion, all attackers are required to provide one another with correct information about their received copies. Then, the colluders adjust the collusion parameters accordingly to ensure that the risk is evenly distributed among all attackers.

Most prior work assumed that all colluders keep their fair collusion agreement and give each other correct information about their fingerprinted signals. During collusion, colluders use the copies that they received from the content owner. In reality, some colluders might be selfish and break their fair-play agreement. They still wish to participate in and profit from collusion, while they do not want to take any risk of being detected by the digital rights enforcer. To achieve this goal, they may lie to other attackers about their fingerprinted copies. For example, they may process their fingerprinted signals before multiuser collusion and use the processed copies instead of the originally received ones during collusion. The selfish colluders' goal is to minimize their own risk while still profiting from collusion. Therefore, during precollusion processing, the selfish colluders select the most effective strategy to reduce their risk. Meanwhile, in order to profit from collusion, the selfish colluders wish that others cannot detect their precollusion processing behavior and will not exclude them from collusion. This requires that the processed copy be perceptually similar to the originally received one and puts stringent quality constraints on precollusion processing.[1]

Depending on the precollusion processing strategies as well as the total number of selfish colluders, in some scenarios, precollusion can increase the absolute risk of other attackers (i.e., their probability of being detected) and it is not only selfish but also malicious. In other scenarios, precollusion processing may have a negligible impact on other colluders' probability of being caught. Another possibility is that precollusion processing decreases the other colluders' absolute risk. Nevertheless, precollusion processing reduces the selfish colluders' risk, makes other attackers have a higher probability of being detected than the selfish colluders and, therefore, increases the relative risk taken by other attackers when compared with that of the selfish colluders. Meanwhile, other attackers are not aware of such precollusion processing and the increase in their relative risk. It is obviously a selfish behavior.

With the existence of selfish colluders, attackers do not trust each other and this distrust forbids them to collude. To continue the collusion attack, the colluders must share something in common that enables them to detect and identify selfish colluders and exclude them from collusion, force everyone to keep their agreement during collusion, and establish the trust among themselves. If all colluders process their received copies before collusion and each acts individually, it corresponds to the scenario where a colluder trusts no one but himself or herself. As such, it is impossible to establish the trust among attackers that are required to continue the collusion attack, and such a group of attackers cannot collude. In this paper, we consider the scenario where most colluders keep their agreement of sharing the risk with others and there are only a few selfish colluders who might process their fingerprinted copies before collusion. In this scenario, those attackers who keep their agreement can collaborate with each other to detect and identify selfish colluders.

This paper studies this problem of "traitors within traitors" in multimedia forensics and formulates this dynamics among attackers during collusion. As the first work on understanding the colluders' behavior to minimize their own risk and protect their own interest, in this paper, we illustrate our framework with a few possible precollusion processing strategies, analyze their performance, and identify the best one for selfish colluders to minimize their risk under quality constraints. Another important issue in this behavior dynamics formulation is to explore the strategies for other attackers to detect such selfish behavior and evaluate their performance. We will investigate this issue in the future.

Game theory is one fundamental tool to formulate the traitor-within-traitor behavior forensics. Such a game-theoretic framework would include the definition of cost functions and the derivation of strategy that maximizes the payoff function. One may ask whether there exists an equilibrium and how the equilibrium strategies can be established, which we plan to explore in the future.

C. Performance Criteria

To measure the effectiveness of precollusion processing in reducing the selfish colluder's probability of being detected, we use the probability that a colluder $\mathbf{u}^{(i)}$ is captured $(P_d^{(i)})$ and the probability that an innocent user is falsely accused (P_{fa}) as the performance criteria. For a fixed P_{fa}, we compare a selfish colluder's probability of being detected in two scenarios: when the selfish colluder does not apply precollusion processing (i.e., he or she is willing to share the risk with other colluders), and when the selfish colluder processes his or her fingerprinted copy before collusion. From the selfish colluder's point of view, precollusion processing is more effective when the difference between these two probabilities is larger.

To measure the effect of precollusion processing on the perceptual quality of the fingerprinted copies, we use the commonly used mean square error (MSE) between the newly generated copy $\tilde{\mathbf{X}}^{(i_1)}$ and the originally received one $\mathbf{X}^{(i_1)}$, or equivalently, the PSNR in image and video applications.

III. Temporal Frame Filtering During Precollusion Processing

For a selfish colluder to further reduce his or her own probability of being detected, one possible solution is to attenuate

[1]In order to avoid being detected by their fellow attackers, selfish colluders should also ensure that the processed copy is statistically similar to the originally received one. We plan to investigate this issue in the future.

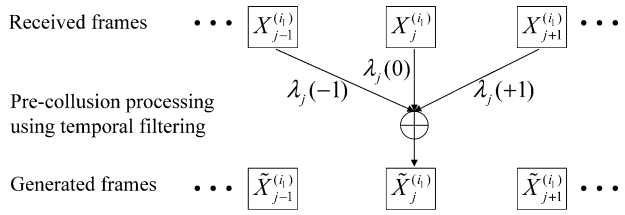

Fig. 1. Applying temporal frame averaging during precollusion processing.

the energy of the embedded fingerprints before multiuser collusion. An example is to replace each segment of the fingerprinted signal with another, a seemingly similar segment from different regions of the content (e.g., averaging or swapping consecutive frames of similar content [18]).

In this section, we take the temporal filtering of adjacent frames as an example, and analyze its effects on the selfish colluder's probability of being detected as well as the perceptual quality of the fingerprinted copies. We consider a simple scenario where all users receive fingerprinted copies of the same quality. When different users receive copies of different quality, the analysis is similar and not repeated.

A. Precollusion Processing Using Temporal Filtering

In this paper, we assume that the selfish colluder uses a simple linear interpolation-based frame average during precollusion processing. A selfish colluder can also apply more complicated motion-based interpolation [21], and the analysis will be similar. For a selfish colluder $\mathbf{u}^{(i_1)}$, assume that $\{\mathbf{X}_j^{(i_1)}\}_{j=1,2,...}$ are the fingerprinted frames that he or she received from the content owner, and $\mathbf{X}_{j-1}^{(i_1)}, \mathbf{X}_j^{(i_1)}$ and $\mathbf{X}_{j+1}^{(i_1)}$ are three consecutive frames. As shown in Fig. 1, for each frame j, $\mathbf{u}^{(i_1)}$ linearly combines the current frame $\mathbf{X}_j^{(i_1)}$, the previous frame $\mathbf{X}_{j-1}^{(i_1)}$, and the next frame $\mathbf{X}_{j+1}^{(i_1)}$ with weights $\lambda_j(0)$, $\lambda_j(-1)$ and $\lambda_j(+1)$, respectively, and generates a new frame $\tilde{\mathbf{X}}_j^{(i_1)}$, where

$$\tilde{\mathbf{X}}_j^{(i_1)} = \lambda_j(-1)\cdot\mathbf{X}_{j-1}^{(i_1)} + \lambda_j(0)\cdot\mathbf{X}_j^{(i_1)} + \lambda_j(+1)\cdot\mathbf{X}_{j+1}^{(i_1)}. \quad (3)$$

In (3), $0 \le \lambda_j(-1), \lambda_j(0), \lambda_j(+1) \le 1$, and $\lambda_j(-1)+\lambda_j(0)+\lambda_j(+1) = 1$. For simplicity, we let $\lambda_j(-1) = \lambda_j(+1) = (1-\lambda_j(0))/2$, and give equal weights to the two neighboring frames $\mathbf{X}_{j-1}^{(i_1)}$ and $\mathbf{X}_{j+1}^{(i_1)}$. $\mathbf{u}^{(i_1)}$ repeats this process for every frame in the sequence and generates $\{\tilde{\mathbf{X}}_j^{(i_1)}\}_{j=1,2,...}$. When $\lambda_j(0) = 1, \tilde{\mathbf{X}}_j^{(i_1)} = \mathbf{X}_j^{(i_1)}$ and it corresponds to the scenario where $\mathbf{u}^{(i_1)}$ does not process his or her copy before collusion.

We assume that there is only one selfish colluder and other colluders do not discover his or her precollusion processing actions. The analysis is similar when there are multiple selfish colluders and is not repeated here. In this scenario, under the averaging collusion, the jth frame in the colluded copy is shown in (4) at the bottom of the page, where $\mathbf{n}_j$ is additive noise.

B. Performance Analysis and Selection of the Optimal Weight Vector

During precollusion processing, the selfish colluder wishes to generate a new copy of high quality and minimize his or her own risk of being detected. In this section, we first analyze the quality of the newly generated frames $\{\tilde{\mathbf{X}}_j^{(i_1)}\}$ and calculate the selfish colluder's probability of being detected, and then study the selection of the optimal weight vector $[\lambda_1(0), \lambda_2(0), \ldots]$.

1) Perceptual Quality: If $\tilde{\mathbf{X}}_j^{(i_1)}$ is generated as in (3), then the MSE between $\tilde{\mathbf{X}}_j^{(i_1)}$ and $\mathbf{X}_j^{(i_1)}$ is

$$\mathrm{MSE}_j = \left\|\tilde{\mathbf{X}}_j^{(i_1)} - \mathbf{X}_j^{(i_1)}\right\|^2 = \left(\frac{1-\lambda_j(0)}{2}\right)^2 \cdot \phi_j$$

where

$$\begin{aligned}\phi_j = {} & 4\left\|\mathbf{X}_j^{(i_1)}\right\|^2 + \left\|\mathbf{X}_{j-1}^{(i_1)}\right\|^2 + \left\|\mathbf{X}_{j+1}^{(i_1)}\right\|^2 \\ & - 4\left\langle\mathbf{X}_{j-1}^{(i_1)}, \mathbf{X}_j^{(i_1)}\right\rangle - 4\left\langle\mathbf{X}_j^{(i_1)}, \mathbf{X}_{j+1}^{(i_1)}\right\rangle \\ & + 2\left\langle\mathbf{X}_{j-1}^{(i_1)}, \mathbf{X}_{j+1}^{(i_1)}\right\rangle. \end{aligned} \quad (5)$$

In (5), $\|\mathbf{X}_j^{(i_1)}\|$ is the Euclidean norm of $\mathbf{X}_j^{(i_1)}$, and $\langle\mathbf{X}_{j-1}^{(i_1)}, \mathbf{X}_j^{(i_1)}\rangle$ is the correlation between $\mathbf{X}_{j-1}^{(i_1)}$ and $\mathbf{X}_j^{(i_1)}$. From (5), a larger $\lambda_j(0)$ implies a smaller MSE_j. Consequently, from the perceptual quality's point of view, $\mathbf{u}^{(i_1)}$ should choose a larger $\lambda_j(0)$. Compared with $\mathbf{X}_j^{(i_1)}, \tilde{\mathbf{X}}_j^{(i_1)}$ has the best possible quality when $\lambda_j(0) = 1$ and $\mathbf{u}^{(i_1)}$ does not apply precollusion processing.

2) Probability of Being Detected: Given the colluded copy $\mathbf{V}'_j$ as in (4), the fingerprint extracted from the jth frame is shown in (6) at the bottom of the page, where $\mathbf{d}_j$ contains terms that are independent of the embedded fingerprints $\{\mathbf{W}_j^{(i)}\}$.

With orthogonal fingerprint modulation as in Section II-A1, given the colluder set SC and the index of the selfish colluder

$$\mathbf{V}'_j = \frac{\sum_{i\in\mathrm{SC}, i\neq i_1}\mathbf{X}_j^{(i)}}{K} + \frac{\lambda_j(-1)\cdot\mathbf{X}_{j-1}^{(i_1)} + \lambda_j(0)\cdot\mathbf{X}_j^{(i_1)} + \lambda_j(+1)\cdot\mathbf{X}_{j+1}^{(i_1)}}{K} + \mathbf{n}_j \quad (4)$$

$$\mathbf{Y}_j = \frac{\sum_{i\in\mathrm{SC}, i\neq i_1}\mathbf{W}_j^{(i_1)}}{K} + \frac{\lambda_j(-1)\cdot\mathbf{W}_{j-1}^{(i_1)} + \lambda_j(0)\cdot\mathbf{W}_j^{(i)} + \lambda_j(+1)\cdot\mathbf{W}_{j+1}^{(i_1)}}{K} + \mathbf{d}_j \quad (6)$$

i_1, if the detection noise $\mathbf{d}_j$ is i.i.d. and follows Gaussian distribution $\mathcal{N}(0,\sigma_n^2)$, it is straightforward to show that the detection statistics $T_N^{(i)}$ in (2) are independent Gaussian with marginal distribution $N(\mu^{(i)},\sigma_n^2)$. The detection statistics have zero mean for an innocent user and positive mean for a guilty colluder. Consequently, given a user $\mathbf{u}^{(i)}$, the probability of accusing him or her if he or she is innocent is $P_{fa}=Q(h/\sigma_n)$, and the probability of capturing him or her if he or she is guilty is $P_d^{(i)}=Q((h-\mu^{(i)})/\sigma_n)$. $Q(\cdot)$ here is the Gaussian tail function and h is the predetermined threshold. For a fixed P_{fa}, the selfish colluder $\mathbf{u}^{(i)}$ has a smaller probability of being detected when $\mu^{(i)}$ is smaller, and minimizing his or her probability of being detected is equivalent to minimizing the mean of his or her detection statistics.

For the selfish colluder $\mathbf{u}^{(i_1)}$, $\mu^{(i_1)}$ is shown in (7) at the bottom of the page. $\|\mathbf{W}_j^{(i_1)}\|$ is the Euclidean norm of $\mathbf{W}_j^{(i_1)}$, $\langle \mathbf{W}_{j-1}^{(i_1)},\mathbf{W}_j^{(i_1)}\rangle$ is the correlation between $\mathbf{W}_{j-1}^{(i_1)}$ and $\mathbf{W}_j^{(i_1)}$, and $\langle \mathbf{W}_j^{(i_1)},\mathbf{W}_{j+1}^{(i_1)}\rangle$ is the correlation between $\mathbf{W}_j^{(i_1)}$ and $\mathbf{W}_{j+1}^{(i_1)}$. From the fingerprint design in Section II-A1, $\langle \mathbf{W}_{j-1}^{(i_1)},\mathbf{W}_j^{(i_1)}\rangle \le \langle \mathbf{W}_j^{(i_1)},\mathbf{W}_j^{(i)}\rangle = \|\mathbf{W}_j^{(i_1)}\|^2$ and $\langle \mathbf{W}_j^{(i_1)},\mathbf{W}_{j+1}^{(i_1)}\rangle \le \|\mathbf{W}_j^{(i_1)}\|^2$. Thus, if $\lambda_1(0),\ldots,\lambda_{j-1}(0),\lambda_{j+1}(0),\ldots$ are fixed, $\mu^{(i_1)}$ is a nondecreasing function of $\lambda_j(0)$ and is minimized when $\lambda_j(0)=0$. Consequently, from the risk minimization's point of view, a smaller $\lambda_j(0)$ is preferred.

3) Selection of the Optimal Weight Vector: From the above analysis, we have seen that during precollusion processing, a selfish colluder should choose larger weights $\{\lambda_j(0)\}$ to minimize the perceptual distortion introduced into his or her fingerprinted copy; while smaller weights $\{\lambda_j(0)\}$ are preferred to minimize his or her risk of being captured. A selfish colluder wishes to minimize his or her probability of being detected while still maintaining good quality of the fingerprinted copies. Thus, for a selfish colluder $\mathbf{u}^{(i_1)}$, the selection of the weight vector $[\lambda_1(0),\lambda_2(0),\ldots]$ can be modeled as

$$\begin{aligned}&\min_{\{\lambda_j(0)\}}\left\{\mu^{(i_1)}=\textstyle\sum_j \mu_j^{(i_1)}\right\}\\ &\text{s.t.}\quad \mathrm{MSE}_j\le\varepsilon, 0\le\lambda_j(0)\le 1,\quad j=1,2,\ldots,\end{aligned}\tag{8}$$

where ε is the constraint on perceptual distortion. In our model of temporal filtering, $\{\lambda_j(0)\}$ for different frames is selected independently. Thus, minimizing $\mu^{(i_1)}$ over the entire video sequence is equivalent to minimizing $\mu_j^{(i_1)}$ in (7) for each frame j independently. Therefore, the optimization problem in (8) is equivalent to: for each frame j

$$\begin{aligned}&\min_{\lambda_j(0)}\mu_j^{i_1}\\ &\text{s.t.}\quad \mathrm{MSE}_j\le\varepsilon,\quad 0\le\lambda_j(0)\le 1.\end{aligned}\tag{9}$$

Given ϕ_j as defined in (5), we can show that the solution to (9) is

$$\lambda_j^*=\max\{0,1-2\sqrt{\varepsilon/\phi_j}\}.\tag{10}$$

By using $\{\lambda_j^*\}$ as in (10) during temporal filtering, a selfish colluder minimizes his or her own probability of being detected and ensures that the newly generated frames have small perceptual distortion ($\mathrm{MSE}\le\varepsilon$) when compared with the originally received ones.

C. Simulation Results

In our simulations, we use the first 40 frames in sequence "carphone" as an example. At the content owner's side, we adopt the human visual model-based spread-spectrum embedding [17], and embed fingerprints in the discrete cosine transform (DCT) domain. We generate independent vectors from Gaussian distribution $\mathcal{N}(0,1/9)$, and then apply Gram–Schmidt orthogonalization to produce fingerprints that satisfy (1) strictly. In each fingerprinted copy, similar to the work in [19] and [20], fingerprints embedded in adjacent frames are correlated with each other, and the correlation depends on the similarity between the two host frames.

At the colluders' side, we assume that there are a total of 150 colluders. For simplicity, we assume that there is only one selfish colluder and he or she applies temporal filtering to his or her received copy as in (3) during precollusion processing. In our simulations, we adjust the power of the noise term $\mathbf{d}_j$ in (6) such that $\|\mathbf{d}_j\|^2=2\|\mathbf{W}_j^{(i)}\|^2$. Other values will give the same trend.

Fig. 2 shows the simulation results. For each frame, j, PSNR_j is defined as the peak signal-to-noise ratio (PSNR) of $\tilde{\mathbf{X}}_j^{(i_1)}$ compared to $\mathbf{X}_j^{(i_1)}$. In Fig. 2, $\{\lambda_j^*\}$ are the solution of (10) and ε is chosen to satisfy $\mathrm{PSNR}_j\ge 40$ dB for all frames. In our simulations, we consider four different scenarios where $\lambda_j(0)=1,\lambda_j(0)=0.8,\lambda_j(0)=\lambda_j^*$, and $\lambda_j(0)=0$, respectively. Note that $\lambda_j(0)=1$ corresponds to the scenario where

$$\mu^{(i_1)}=\sum_j \mu_j^{(i_1)},$$

where

$$\mu_j^{(i_1)}=\frac{\left\langle \mathbf{W}_{j-1}^{(i_1)},\mathbf{W}_j^{(i_1)}\right\rangle+\left\langle \mathbf{W}_j^{(i_1)},\mathbf{W}_{j+1}^{(i_1)}\right\rangle}{2K\sqrt{\sum_l\left\|\mathbf{W}_l^{(i_1)}\right\|^2}}+\lambda_j(0)\times\frac{2\left\|\mathbf{W}_j^{(i_1)}\right\|^2-\left\langle \mathbf{W}_{j-1}^{(i_1)},\mathbf{W}_j^{(i_1)}\right\rangle-\left\langle \mathbf{W}_j^{(i_1)},\mathbf{W}_{j+1}^{(i_1)}\right\rangle}{2\,\mathrm{K}\sqrt{\sum_l\left\|\mathbf{W}_l^{(i_1)}\right\|^2}}\tag{7}$$

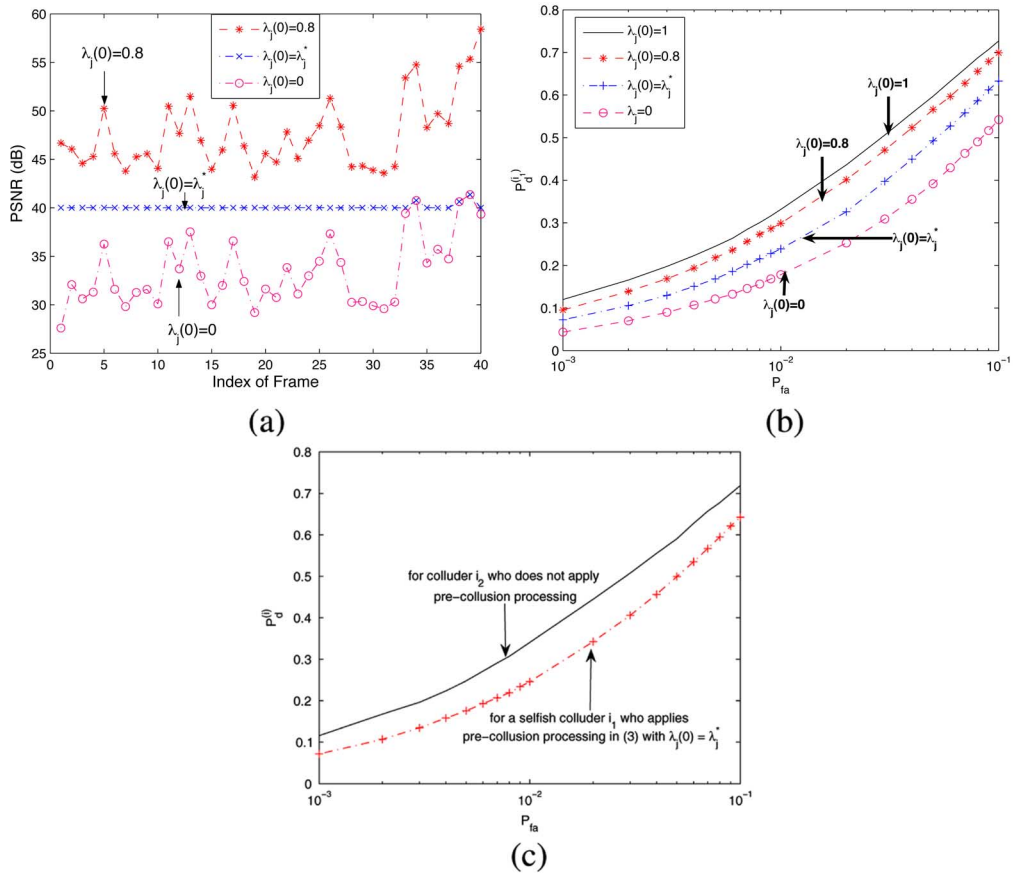

Fig. 2. Simulation results of temporal filtering on sequence "carphone." The length of the embedded fingerprints is 159608. Assume that there are a total of $K = 150$ colluders and there is only one selfish colluder $\mathbf{u}^{(i_1)}$. Colluder $\mathbf{u}^{(i_2)}$ does not process his or her received copy before multiuser collusion. $\{\lambda_j^*\}$ are the solution of (10) and ε is chosen to satisfy $\text{PSNR}_j \geq 40$ dB for all frames. (a) PSNR of the newly generated copy $\{\bar{\mathbf{X}}_j^{(i_1)}\}$ compared with the originally received fingerprinted frames $\{\mathbf{X}_j^{(i_1)}\}$. (b) The selfish colluder's probability of being detected $P_d^{(i_1)}$. (c) Comparison of $\mathbf{u}^{(i_2)}$'s probability of being detected with the selfish colluder $\mathbf{u}^{(i_1)}$'s probability of being detected.

the selfish colluder $\mathbf{u}^{(i_1)}$ does not process his or her copy before multiuser collusion.

Fig. 2(a) compares the perceptual quality of $\{\tilde{\mathbf{X}}_j^{(i_1)}\}$, and Fig. 2(b) plots the selfish colluder $\mathbf{u}^{(i_1)}$'s probability of being detected when $\{\lambda_j(0)\}$ take different values. A selfish colluder can reduce his or her own probability of being detected by temporally filtering his or her fingerprinted copy before multiuser collusion. By choosing $\{\lambda_j(0)\}$ of smaller values, the selfish colluder has a smaller probability of being detected while sacrificing the quality of the newly generated copy. Therefore, during precollusion processing, the selfish colluder has to consider the tradeoff between the risk and the perceptual quality.

In Fig. 2(c), we consider two colluders—the selfish colluder $\mathbf{u}^{(i_1)}$ and another colluder $\mathbf{u}^{(i_2)}$, who does not process his or her copy before collusion, and compare their probabilities of being detected. From Fig. 2(c), precollusion processing makes $\mathbf{u}^{(i_2)}$ take a higher risk of being detected than $\mathbf{u}^{(i_1)}$ and increases the relative risk taken by $\mathbf{u}^{(i_2)}$ when compared with that of $\mathbf{u}^{(i_1)}$.

To address the tradeoff between perceptual quality and the risk, a selfish colluder should choose $\{\lambda_j(0)\}$ as in (10). We compare the solution of $\{\lambda_j^*\}$ in (10) for different sequences. We choose four representative video sequences: "miss america" that has large smooth regions and slow motion, "carphone" and "foreman" that are moderately complicated, and "flower" whose high-frequency band has a lot of energy and the camera moves quickly. We choose the threshold ε in (10) such that $\text{PSNR}_j \geq 40$ dB for all frames in $\{\tilde{\mathbf{X}}_j^{(i_1)}\}$. Fig. 3 shows the solutions of (10) for various sequences. From Fig. 3, for sequences that have slow motion ("miss america"), a selfish colluder can choose $\{\lambda_j(0)\}$ with small values (e.g., around 0), without significant

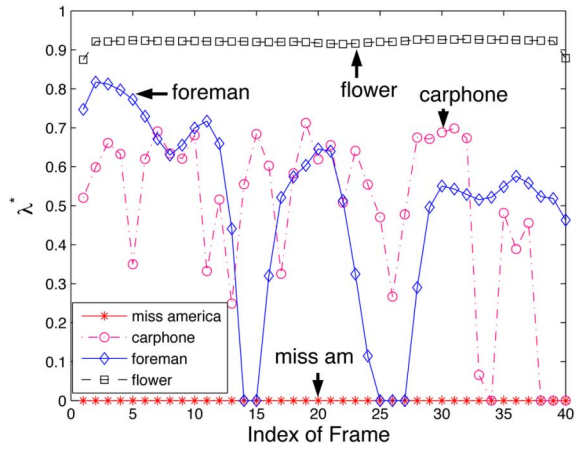

Fig. 3. λ_j^* in (10) for different sequences. ε is chosen to satisfy $\text{PSNR}_j \geq 40$ dB for all frames in $\{\bar{\mathbf{X}}_j^{(i_1)}\}$.

quality degradation; for sequences that have moderate motion ("carphone" and "foreman"), λ_j^* is around 0.5; while for sequences with fast movement ("flower"), a selfish colluder has to choose large $\{\lambda_j(0)\}$ (e.g., larger than 0.9), to ensure the high quality of the newly generated frames.

IV. Traitors Within Traitors in Scalable Fingerprinting Systems

In scalable multimedia coding systems, for the same multimedia content, different users receive copies of different resolutions and quality, depending on each user's available bandwidth and computation constraints. In scalable multimedia coding and fingerprinting systems, in addition to applying temporal filtering to the received frames, a selfish colluder can also change the resolution of his or her fingerprinted copy before multiuser collusion. In this section, we investigate how selfish colluders behave before multiuser collusion in scalable multimedia fingerprinting systems, and analyze their performance.

A. Temporally Scalable Video Coding Systems

Scalable video coding is widely used to accommodate heterogenous networks and users with different computation capability [22]. As an example, we consider temporally scalable video coding, which provides multiple versions of the same video with different temporal resolutions or frame rates. In addition, we use layered video coding to decompose the video sequence into nonoverlapping bit streams of different priority. The base layer contains the most important information of the video content, provides the roughest resolution of the video, and is received by all users in the systems. The enhancement layers contain less important information, gradually refines the reconstructed video at the decoder's side, and are only received by users who have sufficient bandwidth and computation capability.

Without loss of generality, this paper considers a temporally scalable video coding system with three-layer scalability: the base layer of the highest priority, the enhancement layer 1 of medium priority, and the enhancement layer 2 of the lowest priority. Similar to that in [16], a simple implementation of the temporal scalability is used in this paper where different frames are encoded in different layers. For example, with MPEG-2 video coding, the base layer may contain all of the I frames, the enhancement layer 1 contains all of the P frames, and the enhancement layer 2 contains all of the B frames. Assume that F_b, F_{e1} and F_{e2} are the sets containing the indices of the frames that are encoded in the base layer, enhancement layer 1, and enhancement layer 2, respectively.

We let $F^{(i)}$ contain the indices of the frames that user $\mathbf{u}^{(i)}$ receives. Define $\mathbf{U}^b \triangleq \{\mathbf{u}^{(i)} : F^{(i)} = F_b\}$ as the subgroup of users who subscribe to copies of low quality and receive the base-layer bit stream only; $\mathbf{U}^{b,e1} \triangleq \{\mathbf{u}^{(i)} : F^{(i)} = F_b \cup F_{e1}\}$ is the subgroup of users who subscribe to copies of medium quality and receive both the base layer and the enhancement layer 1; and $\mathbf{U}^{\text{all}} \triangleq \{\mathbf{u}^{(i)} : F^{(i)} = F_b \cup F_{e1} \cup F_{e2}\}$ is the subgroup of users who subscribe to copies of high quality and receive all three layers. $\mathbf{U}^b$, $\mathbf{U}^{b,e1}$ and $\mathbf{U}^{\text{all}}$ are mutually exclusive, and $M = |\mathbf{U}^b| + |\mathbf{U}^{b,e1}| + |\mathbf{U}^{\text{all}}|$ is the total number of users.

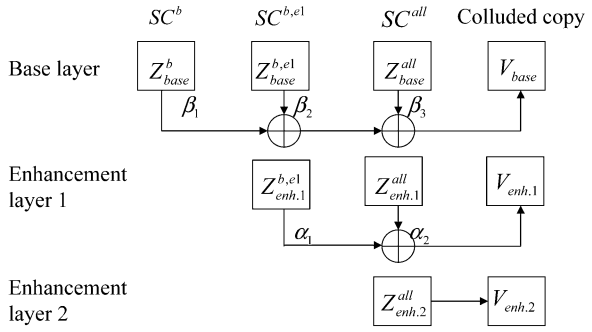

Fig. 4. Two-stage collusion attacks on scalable fingerprinting systems.

B. Scalable Fingerprinting Systems

1) Fingerprint Embedding: With the temporally scalable video coding systems in Section IV-A, the fingerprint embedding at the content owner's side is similar to that in Section II-A1. For each user in the system, and for each frame that he or she subscribes to, the content owner generates a unique fingerprint and additively embeds it into the host signal using spreading-spectrum embedding techniques [17]. Adjacent frames in each distributed copy are embedded with correlated fingerprints to combat intracontent collusion attacks [19], and we consider orthogonal fingerprint modulation in this paper.

2) Collusion Attacks: Assume that there are a total of K colluders and SC is the set containing their indices. We first consider the scenario where all colluders are willing to share the same risk and they provide one another correct information about their fingerprinted copies during collusion. From [16], to generate a colluded copy of high quality while still achieving fairness of collusion, the attackers apply the two-stage collusion, as shown in Fig. 4.

During collusion, the colluders first divide themselves into three nonoverlapping subgroups $\text{SC}^b \triangleq \{i \in \text{SC} : F^{(i)} = F_b\}$ contains the indices of the colluders who receive the base layer bit stream only; $\text{SC}^{b,e1} \triangleq \{i \in \text{SC} : F^{(i)} = F_b \cup F_{e1}\}$ includes the indices of the colluders who receive the base layer and the enhancement layer 1; and $\text{SC}^{\text{all}} \triangleq \{i \in \text{SC} : F^{(i)} = F_b \cup F_{e1} \cup F_{e2}\}$ is the set containing the indices of the colluders who receive all three layers. $K^b, K^{b,e1}$ and K^{all} are the number of colluders in subgroups $\text{SC}^b, \text{SC}^{b,e1}$ and SC^{all}, respectively, and $K^b + K^{b,e1} + K^{\text{all}} = K$.

Secondly, the colluders apply the intragroup collusion attacks: for colluders who receive the base layer only, they average their fingerprinted copies and generate $\mathbf{Z}_j^b = \sum_{i \in \text{SC}^b} \mathbf{X}_j^{(i)} / K^b$ for each frame $j \in F_b$ in the base layer; for colluders who receive both the base layer and the enhancement layer 1, they average their received copies and generate $\mathbf{Z}_j^{b,e1} = \sum_{i \in \text{SC}^{b,e1}} \mathbf{X}_j^{(i)} / K^{b,e1}$ for each frame $j \in F_b \cup F_{e1}$ in the base layer and the enhancement layer 1; and for those colluders who receive all three layers, they average their copies and generate $\mathbf{Z}_j^{\text{all}} = \sum_{i \in \text{SC}^{\text{all}}} \mathbf{X}_j^{(i)} / K^{\text{all}}$ for all frames in the video sequence.

TABLE I
CONSTRAINTS ON COLLUSION TO ACHIEVE FAIRNESS OF THE ATTACK

$F^c = F_b \cup F_{e1} \cup F_{e2}$ (Highest resolution)	$\begin{cases} \frac{K^b\sqrt{N_b}}{K^b\sqrt{N_b}+K^{b,e1}\sqrt{N_b+N_{e1}}+K^{all}\sqrt{N_b+N_{e1}+N_{e2}}} \le \frac{N_b}{N_b+N_{e1}+N_{e2}}, \\ \frac{K^{all}\sqrt{N_b+N_{e1}+N_{e2}}}{K^b\sqrt{N_b}+K^{b,e1}\sqrt{N_b+N_{e1}}+K^{all}\sqrt{N_b+N_{e1}+N_{e2}}} \ge \frac{N_{e2}}{N_b+N_{e1}+N_{e2}}. \end{cases}$
$F^c = F_b \cup F_{e1}$ (Medium resolution)	$\frac{K^b\sqrt{N_b}}{K^b\sqrt{N_b}+(K^{b,e1}+K^{all})\sqrt{N_b+N_{e1}}} \le \frac{N_b}{N_b+N_{e1}}$.
$F^c = F_b$ (Lowest resolution)	No constraints on $(K^b, K^{b,e1}, K^{all})$ and (N_b, N_{e1}, N_{e2}).

TABLE II
SELECTION OF COLLUSION PARAMETERS TO ACHIEVE FAIRNESS OF THE ATTACK

$F^c = F_b \cup F_{e1} \cup F_{e2}$ (Highest resolution)	$\begin{cases} \beta_1 = \frac{N_b+N_{e1}+N_{e2}}{N_b} \frac{K^b\sqrt{N_b}}{K^b\sqrt{N_b}+K^{b,e1}\sqrt{N_b+N_{e1}}+K^{all}\sqrt{N_b+N_{e1}+N_{e2}}}, \\ \beta_2 N_b + \alpha_1 N_{e1} = \frac{(N_b+N_{e1}+N_{e2})K^{b,e1}\sqrt{N_b+N_{e1}}}{K^b\sqrt{N_b}+K^{b,e1}\sqrt{N_b+N_{e1}}+K^{all}\sqrt{N_b+N_{e1}+N_{e2}}}, \\ \beta_3 = 1-\beta_1-\beta_2,\ \alpha_2 = 1-\alpha_1. \end{cases}$
$F^c = F_b \cup F_{e1}$ (Medium resolution)	$\begin{cases} \beta_1 = \frac{N_b+N_{e1}}{N_b} \frac{K_b\sqrt{N_b}}{K^b\sqrt{N_b}+(K^{b,e1}+K^{all})\sqrt{N_b+N_{e1}}}, \\ \beta_2 = \frac{K^{b,e1}}{K^{b,e1}+K^{all}}(1-\beta_1),\ \beta_3 = 1-\beta_1-\beta_2, \\ \alpha_1 = \frac{K^{b,e1}}{K^{b,e1}+K^{all}},\ \alpha_2 = 1-\alpha_1. \end{cases}$
$F^c = F_b$ (Lowest resolution)	$\beta_1 = \frac{K^b}{K^b+K^{b,e1}+K^{all}},\ \beta_2 = \frac{K^{b,e1}}{K^b+K^{b,e1}+K^{all}},\ \beta_3 = \frac{K^{all}}{K^b+K^{b,e1}+K^{all}}$.

Define F^c as the set containing the indices of the frames in the colluded copy. For simplicity, $F^c \in \{F_b,\ F_b \cup F_{e1},\ F_b \cup F_{e1} \cup F_{e2}\}$ and they correspond to the scenarios where the colluded copy has the lowest, medium, and highest frame rates, respectively. Finally, the colluders apply the intergroup collusion attacks to generate the colluded copy $\{\mathbf{V}_j\}$. For each frame $j \in F_b$ in the base layer, $\mathbf{V}_j = \beta_1\mathbf{Z}_j^b + \beta_2\mathbf{Z}_j^{b,e1} + \beta_3\mathbf{Z}_j^{\text{all}} + \mathbf{n}_j$ where $0 \le \beta_1,\ \beta_2,\ \beta_3 \le 1$ and $\beta_1+\beta_2+\beta_3 = 1$. For each frame $j \in F_{e1}$ in the enhancement layer 1, $\mathbf{V}_j = \alpha_1\mathbf{Z}_j^{b,e1} + \alpha_2\mathbf{Z}_j^{\text{all}} + \mathbf{n}_j$, where $0 \le \alpha_1,\ \alpha_2 \le \alpha_1+\alpha_2 = 1$. For each frame $j \in F_{e2}$ in enhancement layer 2, $\mathbf{V}_j = \mathbf{Z}_j^{\text{all}} + \mathbf{n}_j$. $\mathbf{n}_j$ is additive noise to further hinder detection.

3) Fingerprint Detection and Colluder Identification: Same as in [16], this paper considers a simple detector that uses fingerprints extracted from all layers collectively to identify colluders. For each user $\mathbf{u}^{(i)}$, the detector first calculates $\breve{F}^{(i)} \triangleq F^{(i)} \cap F^c$, where $F^{(i)}$ contains the indices of the frames received by user $\mathbf{u}^{(i)}$, and F^c contains the indices of the frames in the colluded copy. Following the thresholding detection strategy in [15], after extracting the fingerprint $\mathbf{Y}_j$ from the colluded frame $\mathbf{V}_j$, the detector calculates $T_N^{(i)} = (\sum_{j\in\breve{F}^{(i)}} \langle \mathbf{Y}_j, \mathbf{W}_j^{(i)} \rangle) / \sqrt{\sum_{l\in\breve{F}^{(i)}} \|\mathbf{W}_l^{(i)}\|^2}$. Given a predetermined threshold h, the estimated colluder set is $\hat{\text{SC}} = \{i : T_N^{(i)} > h\}$.

4) Constraints on Multiuser Collusion to Achieve Fairness: Under the assumption that all colluders are willing to share the same risk and give each other correct information about their received copies, the colluders choose the collusion parameters, including F^c, $\{\beta_k\}_{k=1,2,3}$ and $\{\alpha_l\}_{l=1,2}$ to ensure that all colluders are equally likely to be detected. Given the simple detector in Section IV-B3, Tables I and II list the constraints on collusion and the selection of the collusion parameters, respectively, to achieve fairness in three different scenarios, where the colluded copy $\{\mathbf{V}_j\}$ has the highest, medium, and lowest frame rates, respectively. Detailed derivation is available in [16] and not repeated here. In Tables I and II, N_b, N_{e1}, and N_{e2} are the lengths of the fingerprints embedded in the base layer, enhancement layer 1, and enhancement layer 2, respectively. From Table I, generating a colluded copy of higher quality puts more severe constraints on collusion to achieve fairness.

C. Changing the Resolution of the Fingerprinted Copies Before Collusion

Assume that $F^{(i_1)}$ contains the indices of the frames that a selfish colluder $\mathbf{u}^{(i_1)}$ subscribed to, and $\{\mathbf{X}_j^{(i)}\}_{j\in F^{(i_1)}}$ are the fingerprinted frames that he or she received from the content owner. Before collusion, $\mathbf{u}^{(i_1)}$ processes his or her received copy and generates another copy $\{\tilde{\mathbf{X}}_j^{(i_1)}\}$, whose temporal resolution is different from that of $\{\mathbf{X}_j^{(i_1)}\}$. Assume that $\tilde{F}^{(i_1)}$ contains the indices of the frames in $\{\tilde{\mathbf{X}}_j^{(i_1)}\}$ and $\tilde{F}^{(i_1)} \ne F^{(i_1)}$. During collusion, $\mathbf{u}^{(i_1)}$ uses the newly generated copy $\{\tilde{\mathbf{X}}_j^{(i_1)}\}_{j\in\tilde{F}^{(i_1)}}$, instead of $\{\mathbf{X}_j^{(i_1)}\}_{j\in F^{(i_1)}}$. For simplicity, in this section, we assume that the selfish colluders only change the resolution of their received copies and do not further apply temporal filtering during precollusion processing.

We consider a simple scenario where $\tilde{F}^{(i_1)} \in \{F_b,\ F_b \cup F_{e1},\ F_b \cup F_{e1} \cup F_{e2}\}$. We assume that there is only one selfish colluder $\mathbf{u}^{(i_1)}$ who changes the frame rate of his or her copy before multiuser collusion, and our analysis can be extended to complicated scenarios where there are multiple selfish colluders.

For a selfish colluder $\mathbf{u}^{(i_1)}$ who changes the temporal resolution of his or her copy during precollusion processing, we define the processing parameter as $CP^{(i_1)} \triangleq (F^{(i_1)}, \tilde{F}^{(i_1)})$, where $F^{(i_1)}$ contains the indices of the frames that $\mathbf{u}^{(i_1)}$ received from the content owner and $\tilde{F}^{(i_1)}$ contains the indices of the frames in the newly generated copy $\{\tilde{\mathbf{X}}_j^{(i_1)}\}$. If $\tilde{F}^{(i_1)} \supset F^{(i_1)}$, the selfish

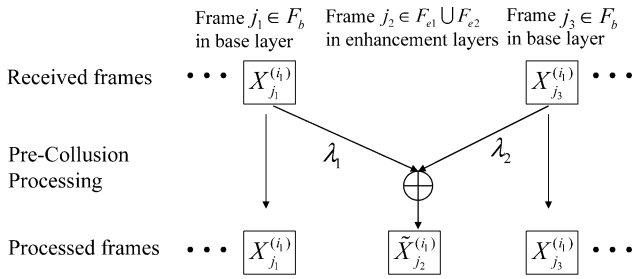

Fig. 5. Example of increasing the temporal resolution during precollusion processing. $F^{(i_1)} = F_b$ and $\tilde{F}^{(i_1)} = F_b \cup F_{e1} \cup F_{e2}$.

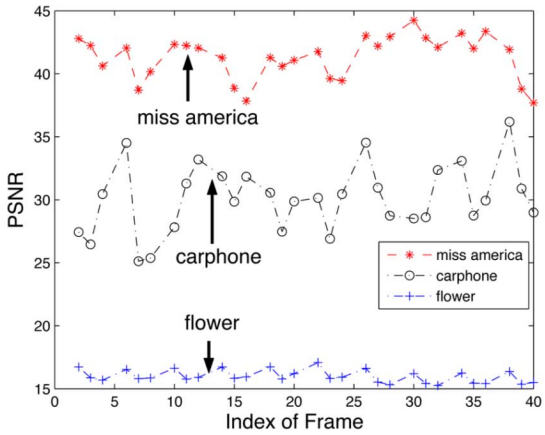

Fig. 6. Quality of the enhancement layers that are forged by the selfish colluder during precollusion processing. The processing parameter is $\mathrm{CP}^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$, where $F_b = \{1, 5, 9, \ldots\}$, $F_{e1} = \{3, 7, 11, \ldots\}$ and $F_{e2} = \{2, 4, 6, 8, \ldots\}$.

colluder $\mathbf{u}^{(i_1)}$ subscribes to a lower quality version and he or she increases the frame rate during precollusion processing. If $\tilde{F}^{(i_1)} \subset F^{(i_1)}$, the selfish colluder $\mathbf{u}^{(i_1)}$ subscribes to a higher quality version and he or she reduces the temporal resolution before multiuser collusion.

1) Increasing the Resolution Before Multiuser Collusion: In this type of precollusion processing, a selfish colluder $\mathbf{u}^{(i_1)}$ subscribes to a copy of a lower frame rate and generates a copy of higher resolution before collusion. Without loss of generality, in this section, we consider the example in Fig. 5 where the processing parameter is $\mathrm{CP}^{(i_1)} = (F^{(i_1)} = F_b, \tilde{F}^{(i_1)} = F_b \cup F_{e1} \cup F_{e2})$. In this example, the selfish colluder receives the fingerprinted base layer only, and generates a copy $\{\tilde{\mathbf{X}}_j^{(i_1)}\}$ with all three layers before collusion. He or she then tells the other colluders that $\{\tilde{\mathbf{X}}_j^{(i_1)}\}_{j \in F_b \cup F_{e1} \cup F_{e2}}$ is the copy that he or she received.

a) Precollusion processing of the fingerprinted copy: We assume that for every frame $j \in F^{(i_1)} = F_b$ in the base layer that $\mathbf{u}^{(i_1)}$ received, the selfish colluder simply duplicates $\mathbf{X}_j^{(i_1)}$ in the newly generated copy and let $\tilde{\mathbf{X}}_j^{(i_1)} = \mathbf{X}_j^{(i_1)}$. $\mathbf{u}^{(i_1)}$ also needs to forge frames $\tilde{\mathbf{X}}_{j \in F_{e1} \cup F_{e2}}^{(i_1)}$ in the enhancement layers that he or she did not receive. Assume that $\mathbf{X}_{j_1}^{(i_1)}$ and $\mathbf{X}_{j_3}^{(i_1)}$ are two adjacent frames in the base layer that $\mathbf{u}^{(i_1)}$ received. To forge a frame $j_2 \in F_{e1} \cup F_{e2}$ in the enhancement layers where $j_1 < j_2 < j_3$, we consider a simple linear interpolation-based method and let $\tilde{\mathbf{X}}_{j_2}^{(i_1)} = \lambda_1 \cdot \mathbf{X}_{j_1}^{(i_1)} + \lambda_2 \cdot \mathbf{X}_{j_3}^{(i_1)}$, where $\lambda_1 = (j_3 - j_2)/(j_3 - j_1)$ and $\lambda_2 = (j_2 - j_1)/(j_3 - j_1)$. Other complicated algorithms (e.g., motion-based interpolation [21]) can be used to improve the quality of the forged frames, and the analysis will be similar.

b) Perceptual quality constraints: To increase the frame rate of the fingerprinted copy, the selfish colluder has to generate frames in the enhancement layers that he or she did not receive from the content owner. To cover up the fact he or she processed the copy before collusion and make other colluders believe him or her, the selfish colluder must ensure that the forged enhancement layers have high quality.

In this section, we examine the perceptual quality of the forged enhancement layers and study the quality constraints. We consider the example in Fig. 5 with processing parameter $\mathrm{CP}^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$, and use the above linear interpolation-based method.

For a selfish colluder $\mathbf{u}^{(i_1)}$ in subgroup SC^b and for a frame $j \in F_{e1} \cup F_{e2}$ in the enhancement layers, define $\mathbf{X}_j^{(i_1)}$ as the fingerprinted frame j that $\mathbf{u}^{(i_1)}$ would have received if he or she had subscribed to frame j. In our simulations, we choose $\mathbf{X}_j^{(i_1)}$ as the ground truth and use the PSNR of $\tilde{\mathbf{X}}_j^{(i_1)}$ when compared with $\mathbf{X}_j^{(i_1)}$ to measure the perceptual quality of the forged frames in the enhancement layers.

Fig. 6 shows the results on the first 40 frames of sequence "miss america," "carphone," and "flower." From Fig. 6, for sequence "miss america" with flat regions and slow motion, the selfish colluder can forge enhancement layers of high quality. For sequence "flower" that has fast movement, the selfish colluder can only generate low-quality and blurred enhancement layers. Therefore, due to the quality constraints, for complicated sequences with fast movement, the selfish colluder might not be able to apply this type of precollusion processing and increase the temporal resolution before multiuser collusion.[2]

c) Selfish colluder's probability of being detected: To analyze the effectiveness of this precollusion processing in reducing a selfish colluder's risk, we compare his or her probability of being detected when the selfish colluder increases the temporal resolution with that when the selfish colluder does not process his or her fingerprinted copy before collusion. Without loss of generality, we assume that the selfish colluder processes his or her copy as in Fig. 5 with parameter $\mathrm{CP}^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$, and use this example to analyze the impact of resolution change on the selfish colluder's probability of being detected.

Scenario 1: Without Precollusion Processing: We first consider the scenario where $\mathbf{u}^{(i_1)}$ does not apply precollusion processing, and we assume that $\mathrm{SC}^b = \{i \in \mathrm{SC} : F^{(i)} = F_b\}$ contains the indices of the colluders who subscribe to copies of the lowest resolution and only receive the base layer from the content owner; $\mathrm{SC}^{b,e1} = \{i \in \mathrm{SC} : F^{(i)} = F_b \cup F_{e1}\}$ contains the indices of the colluders who receive both the base layer and the enhancement layer 1 from the content owner; and

[2]Motion-based interpolation [21] can be used to improve the quality. However, for some sequences with fast movement and complex scene composition (e.g., "football" and "flower"), even with motion-based interpolation, the selfish colluder still may not be able to forge enhancement layers of good enough quality to use. Therefore, for those complicated sequences, the selfish colluders may not be able to increase the resolution of their fingerprinted copies before multiuser collusion.

$\mathrm{SC}^{\mathrm{all}} = \{i \in \mathrm{SC} : F^{(i)} = F_b \cup F_{e1} \cup F_{e2}\}$ contains the indices of the colluders who receive all three layers from the content owner. K^b, $K^{b,e1}$ and K^{all} are the number of colluders in SC^b, $\mathrm{SC}^{b,e1}$ and $\mathrm{SC}^{\mathrm{all}}$, respectively.

Given $(K^b, K^{b,e1}, K^{\mathrm{all}})$ and (N_b, N_{e1}, N_{e2}), the colluders first check the constraints in Table I, and then choose the collusion parameters $\{\beta_k\}_{k=1,2,3}$ and $\{\alpha_l\}_{l=1,2}$ according to Table II. In this scenario, for each frame $j \in F_b$ in the base layer, the extracted fingerprint is

$$\mathbf{Y}_j = \frac{\beta_1 \cdot \mathbf{W}_j^{(i_1)}}{K^b} + \sum_{i \in \mathrm{SC}^b, i \neq i_1} \frac{\beta_1 \cdot \mathbf{W}_j^{(i)}}{K^b} + \sum_{i \in \mathrm{SC}^{b,e1}} \frac{\beta_2 \cdot \mathbf{W}_j^{(i)}}{K^{b,e1}} + \sum_{i \in \mathrm{SC}^{\mathrm{all}}} \frac{\beta_3 \cdot \mathbf{W}_j^{(i)}}{K^{\mathrm{all}}} + \mathbf{n}_j \quad (11)$$

where $\mathbf{n}_j$ is additive noise.

Following the detection procedure in Section IV-B3, the detector observes that $\mathbf{u}^{(i_1)}$ only received the fingerprinted base layer from the content owner and, therefore, the detector will only use fingerprints extracted from the base layer to decide if $\mathbf{u}^{(i_1)}$ is involved in collusion. The detector calculates $T_N^{(i_1)} = (\sum_{j \in F_b} \langle \mathbf{Y}_j, \mathbf{W}_j^{(i_1)} \rangle) / \sqrt{\sum_{l \in F_b} \|\mathbf{W}_l^{(i_1)}\|^2}$, compares it with the predetermined threshold h, and decides if $\mathbf{u}^{(i_1)}$ is a colluder. From the analysis in [16], with orthogonal modulation, given the colluder set SC and the extracted fingerprint as in (11), if the detection noise $\mathbf{n}_j$ is i.i.d. Gaussian $\mathcal{N}(0, \sigma_n^2)$, the detection statistics follow distribution:

$$p\left(T_N^{(i_1)} \,\middle|\, \mathrm{SC}\right) \sim \mathcal{N}\left(\mu^{(i_1)}, \sigma_n^2\right)$$

where

$$\mu^{(i_1)} = \frac{\beta_1}{K^b} \sqrt{\sum_{j \in F_b} \left\|\mathbf{W}_j^{(i_1)}\right\|^2} = \frac{\beta_1 \sqrt{N_b}}{K^b} \xi \quad (12)$$

where $\mathbf{u}^{(i_1)}$'s probability of being detected is $P_d^{(i_1)} = Q((h - \mu^{(i_1)})/(\sigma_n))$, where $Q(\cdot)$ is the Gaussian tail function. In this scenario, all colluders share the same risk and their probability of being detected is equal to $P_d^{(i_1)}$.

Scenario 2: With Precollusion Processing: We then consider the scenario where $\mathbf{u}^{(i_1)}$ increases the frame rate before multiuser collusion and assume that $\widetilde{\mathrm{SC}}^b = \{i \in \mathrm{SC} : \tilde{F}^{(i)}\}$ contains the indices of the colluders who tell others that they received the base layer only; $\widetilde{\mathrm{SC}}^{b,e1} = \{i \in \mathrm{SC} : \tilde{F}^{(i)} = F_b \cup F_{e1}\}$ is the set containing the indices of the colluders who tell others that they received both the base layer and enhancement layer 1; and $\widetilde{\mathrm{SC}}^{\mathrm{all}} = \{i \in \mathrm{SC} : \tilde{F}^{(i)} = F_b \cup F_{e1} \cup F_{e2}\}$ is the set containing the indices of the colluders who tell others that they received all three layers. Define $\tilde{K}^b$, $\tilde{K}^{b,e1}$ and $\tilde{K}^{\mathrm{all}}$ as the number of colluders in $\widetilde{\mathrm{SC}}^b$, $\widetilde{\mathrm{SC}}^{b,e1}$ and $\widetilde{\mathrm{SC}}^{\mathrm{all}}$, respectively.

If $\mathbf{u}^{(i_1)}$ is the only selfish colluder and the processing parameter is $\mathrm{CP}^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$, then we have $\widetilde{\mathrm{SC}}^b = \mathrm{SC}^b \setminus \{i_1\}$, $\widetilde{\mathrm{SC}}^{b,e1} = \mathrm{SC}^{b,e1}$ and $\widetilde{\mathrm{SC}}^{\mathrm{all}} = \mathrm{SC}^{\mathrm{all}} \cup \{i_1\}$. Consequently, $\tilde{K}^b = K^b - 1$, $\tilde{K}^{b,e1} = K^{b,e1}$ and $\tilde{K}^{\mathrm{all}} = K^{\mathrm{all}} + 1$. If other colluders do not discover $\mathbf{u}^{(i_1)}$'s precollusion processing, they assume that the extracted fingerprints from all threes layers will be used by the detector to determine whether $\mathbf{u}^{(i_1)}$ is a colluder. Under this assumption, the colluders analyze each attacker's detection statistics and follow Table II to choose the collusion parameters.

As an example, assume that the colluders decide to generate a colluded copy including all frames in the base layer and the enhancement layer 1, and $(\tilde{K}^b, \tilde{K}^{b,e1}, \tilde{K}^{\mathrm{all}})$ and (N_b, N_{e1}, N_{e2}) satisfy the constraint $(\tilde{K}^b \sqrt{N_b})/(\tilde{K}^b \sqrt{N_b} + (\tilde{K}^{b,e1} + \tilde{K}^{\mathrm{all}}) \sqrt{N_b + N_{e1}}) \leq (N_b)/(N_b + N_{e1})$ listed in Table I. Following the analysis in [16], under the assumption that fingerprints extracted from both layers would be used by the detector to identify $\mathbf{u}^{(i_1)}$, other colluders estimate that $\mathbf{u}^{(i_1)}$'s detection statistics have mean

$$\bar{\mu}^{(i_1)} = \frac{\tilde{\beta}_3 N_b + \tilde{\alpha}_2 N_{e1}}{\tilde{K}^{\mathrm{all}} \sqrt{N_b + N_{e1}}} \xi \quad (13)$$

where N_b and N_{e1} are the lengths of the fingerprints embedded in the base layer and the enhancement layer 1, respectively. They choose the collusion parameters such that $\bar{\mu}^{(i_1)}$ is equal to the means of other colluders' detection statistics. From Table II, the selected parameters are

$$\begin{aligned} \tilde{\beta}_1 &= \frac{N_b + N_{e1}}{N_b} \cdot \frac{\tilde{K}^b \sqrt{N_b}}{\tilde{K}^b \sqrt{N_b} + (\tilde{K}^{b,e1} + \tilde{K}^{\mathrm{all}}) \sqrt{N_b + N_{e1}}} \\ \tilde{\beta}_2 &= \frac{\tilde{K}^{b,e1}}{\tilde{K}^{b,e1} + \tilde{K}^{\mathrm{all}}} \left(1 - \tilde{\beta}_1\right), \quad \tilde{\beta}_3 = 1 - \tilde{\beta}_1 - \tilde{\beta}_2 \\ \tilde{\alpha}_1 &= \frac{\tilde{K}^{b,e1}}{\tilde{K}^{b,e1} + \tilde{K}^{\mathrm{all}}}, \text{ and } \tilde{\alpha}_2 = 1 - \tilde{\alpha}_1. \end{aligned} \quad (14)$$

Then, the colluders generate the colluded copy as in Section IV-B2.

During the colluder identification process, since $\mathbf{u}^{(i_1)}$ only received the fingerprinted base layer from the content owner, the detector only uses fingerprints extracted from the base layer to decide if $\mathbf{u}^{(i_1)}$ is a colluder. The extracted fingerprint from frame $j \in F_b$ in the base layer is

$$\mathbf{Y}_j = \frac{\tilde{\beta}_3 \cdot \mathbf{W}_j^{(i_1)}}{\tilde{K}^{\mathrm{all}}} + \sum_{i \in \mathrm{SC}^b, i \neq i_1} \frac{\tilde{\beta}_1 \cdot \mathbf{W}_j^{(i)}}{\tilde{K}^b} + \sum_{i \in \mathrm{SC}^{b,e1}} \frac{\tilde{\beta}_2 \cdot \mathbf{W}_j^{(i)}}{\tilde{K}^{b,e1}} + \sum_{i \in \mathrm{SC}^{\mathrm{all}}} \frac{\tilde{\beta}_3 \cdot \mathbf{W}_j^{(i)}}{\tilde{K}^{\mathrm{all}}} + \mathbf{n}_j. \quad (15)$$

With orthogonal fingerprint modulation, given the colluder set SC, the index of the selfish colluder i_1, and the precollusion processing parameter $\mathrm{CP}^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$, if $\mathbf{n}_j$ follows Gaussian distribution $\mathcal{N}(0, \sigma_n^2)$ and using the same analysis as in [16], we can show that

$$p\left(T_N^{(i_1)} \,\middle|\, \mathrm{SC}, i_1, \mathrm{CP}^{(i_1)}\right) \sim \mathcal{N}\left(\tilde{\mu}^{(i_1)}, \sigma_n^2\right)$$

and

$$\tilde{P}_d^{(i_1)} = Q\left(\frac{h - \tilde{\mu}^{(i_1)}}{\sigma_n}\right),$$

where

$$\tilde{\mu}^{(i_1)} = \frac{\tilde{\beta}_3}{\tilde{K}^{\mathrm{all}}} \sqrt{\sum_{j \in F_b} \left\|\mathbf{W}_j^{(i_1)}\right\|^2} = \frac{\tilde{\beta}_3 \sqrt{N_b}}{\tilde{K}^{\mathrm{all}}} \xi. \quad (16)$$

For colluder $\mathbf{u}^{(i_2)}$ who does not process his or her copy before collusion, following the same analysis, we can show that his or her detection statistics follow Gaussian distribution $p(T_N^{(i_2)}|\mathrm{SC}, i_1, \mathrm{CP}^{(i_1)}) \sim \mathcal{N}(\tilde{\mu}^{(i_2)}, \sigma_n^2)$, where $\tilde{\mu}^{(i_2)} = \tilde{\beta}_1\sqrt{N_b}\xi/\tilde{K}^b$, and his or her probability of being detected is $P_d^{(i_2)} = Q((h - \tilde{\mu}^{(i_2)})/\sigma_n)$.

Note that $\bar{\mu}^{(i_1)}$ in (13) does not equal $\tilde{\mu}^{(i_1)}$ in (16), and the colluders make an error in estimating the mean of $\mathbf{u}^{(i_1)}$'s detection statistics. This is due to $\mathbf{u}^{(i_1)}$'s precollusion processing behavior, and this estimation error helps the selfish colluder further lower his or her risk of being detected.

From (12) and (16), for fixed h and σ_n^2, comparing the selfish colluder's probability of being detected in these two scenarios is equivalent to comparing $\mu^{(i_1)}$ in (12) with $\tilde{\mu}^{(i_1)}$ in (16). For fair comparison, if the constraints in Table I are satisfied, we fix the frame rate of the colluded copy and let $\tilde{F}^c = F^c$.

To compare the values of the two means, we consider the following scalable fingerprinting systems. We observe that for typical video sequences such as "miss america," "carphone," and "foreman," each frame has approximately 3000–7000 embeddable coefficients, depending on the characteristics of the sequences. As an example, we assume that the length of the embedded fingerprints in each frame is 5000, and we test on a total of 40 frames. We choose $F_b = \{1, 5, 9, \ldots\}$, $F_{e1} = \{3, 7, 11, \ldots\}$ and $F_{e2} = \{2, 4, 8, \ldots\}$ as an example of the temporal scalability, and the lengths of the fingerprints embedded in the base layer, enhancement layer 1 and enhancement layer 2 are $N_b = 50000, N_{e1} = 50000$ and $N_{e2} = 100000$, respectively. We assume that there are a total of $M = 450$ users and $|\mathbf{U}^b| = |\mathbf{U}^{b,e1}| = |\mathbf{U}^{\mathrm{all}}| = 150$. We first generate a unique vector following Gaussian distribution $\mathcal{N}(0, 1/9)$ for each user, and then apply Gram–Schmidt orthogonalization to ensure that the assigned fingerprints satisfy (1) strictly.

We assume that there are a total of $K = 150$ colluders, and $(K^b, K^{b,e1}, K^{\mathrm{all}})$ are on the line as shown in (17) at the bottom of the page. Line (17) is the boundary of one of the constraints in Table I to achieve fairness of collusion when generating a colluded copy of the highest resolution. Other values of $(K^b, K^{b,e1}, K^{\mathrm{all}})$ and (N_b, N_{e1}, N_{e2}) give the same trend.

Assume that there is only one selfish colluder $\mathbf{u}^{(i_1)}$ and $\mathrm{CP}^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$. Fig. 7 compares $\mu^{(i_1)}$ in (12) with $\tilde{\mu}^{(i_1)}$ in (16) when $K = 150$ and $(K^b, K^{b,e1}, K^{\mathrm{all}})$ take different values on line (17). In Fig. 7, a given value of K^b corresponds to a unique point on line (17) and, therefore, a unique triplet $(K^b, K^{b,e1}, K^{\mathrm{all}})$. In Fig. 7(a), $F^c = F_b \cup F_{e1} \cup F_{e2}$, and the colluded copy has the highest resolution; and in Fig. 7(b), $F^c = F_b$, and the colluded copy only contains frames in the base layer. From Fig. 7, increasing the resolution of his or her fingerprinted copy before multiuser collusion can help the selfish colluder further reduce his or her probability of being detected when the colluded copy is of high quality; while it cannot lower the selfish colluder's risk when the colluders decide to generate a copy of the lowest frame rate. This is because when $F^c = F_b$, no matter how many frames that $\mathbf{u}^{(i_1)}$ claims that he or she has received, only those in the base layer are used to generate the colluded copy, and those frames are the ones that $\mathbf{u}^{(i_1)}$ received from the content owner. In this scenario, other colluders correctly estimate the mean of $\mathbf{u}^{(i_1)}$'s detection statistics during collusion, and increasing the frame rate cannot

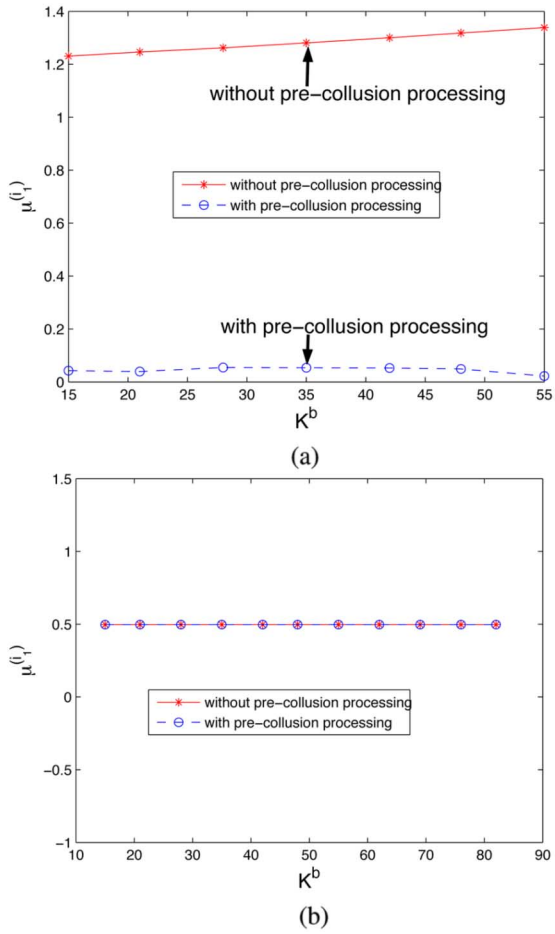

Fig. 7. Comparison of $\mu^{(i_1)}$ in (12) and $\bar{\mu}^{(i_1)}$ in (16) when $(K^b, K^{b,e1}, K^{\mathrm{all}})$ takes different values on line (17). (a) $F^c = F_b \cup F_{e1} \cup F_{e2}$. (b) $F^c = F_b$. $(|F_b|, |F_{e1}|, |F_{e2}|) = (10, 10, 20)$, and $(N_b, N_{e1}, N_{e2}) = (50000, 50000, 100000)$. There are $M = 450$ users and a total of $K = 150$ colluders. Assume that there is only one selfish colluder and the processing parameter is $\mathrm{CP}^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$. Each K^b on the x axis corresponds to a unique point on line (17).

$$\left\{(K^b, K^{b,e1}, K^{\mathrm{all}}) : \frac{K^{\mathrm{all}}\sqrt{N_b + N_{e1} + N_{e2}}}{K^b\sqrt{N_b} + K^{b,e1}\sqrt{N_b + N_{e1}} + K^{\mathrm{all}}\sqrt{N_b + N_{e1} + N_{e2}}} = \frac{N_{e2}}{N_b + N_{e1} + N_{e2}}\right.$$
$$K^b + K^{b,e1} + K^{\mathrm{all}} = 150,$$
$$\left.0 \le K^b \le |\mathbf{U}^b|,\ 0 \le K^{b,e1} \le |\mathbf{U}^{b,e1}|,\ 0 \le K^{\mathrm{all}} \le |\mathbf{U}^{\mathrm{all}}|\right\} \quad (17)$$

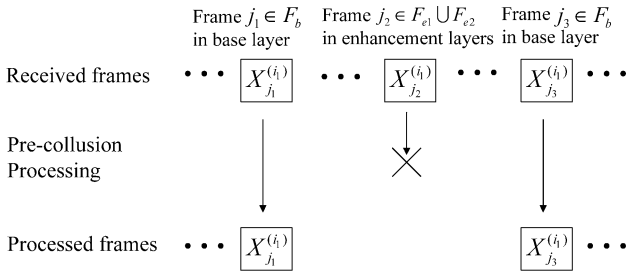

Fig. 8. Example of reducing the frame rate before multiuser collusion. $F^{(i_1)} = F_b \cup F_{e1} \cup F_{e2}$ and $\tilde{F}^{(i_1)} = F_b$.

help the selfish colluder further reduce his or her risk. To generalize, increasing the temporal resolution is effective in reducing a selfish colluder $\mathbf{u}^{(i_1)}$'s probability of being captured only if $F^c \supset F^{(i_1)}$.

2) Reducing the Resolution Before Multiuser Collusion: In this type of precollusion processing, a selfish colluder receives a copy of higher resolution and tells other colluders that he or she only has a copy of lower quality. Shown in Fig. 8 is an example, where $\mathbf{u}^{(i_1)}$ subscribes to all three layers while claiming that he or she only has the fingerprinted base layer. In this example, $\mathbf{u}^{(i_1)}$ simply drops frames in the two enhancement layers during precollusion processing.

When reducing the frame rate of his or her fingerprinted copy, the selfish colluder does not need to forge any frames and, therefore, he or she does not need to worry about the quality constraints. In this scenario, the analysis of the selfish colluder's risk of being detected is similar to that in Section IV-C1 and omitted.

Fig. 9 compares the means of the selfish colluder's detection statistics when he or she drops frames in the enhancement layers with that when he or she does not apply precollusion processing. The setup of the scalable fingerprinting system in Fig. 9 is the same as that in Fig. 7. Similarly, each K^b in Fig. 9 represents one point on Line (17) and a unique $(K^b, K^{b,e1}, K^{\text{all}})$ triplet. The precollusion processing parameter is $\text{CP}^{(i_1)} = (F_b \cup F_{e1} \cup F_{e2}, F_b)$. $F^c = F_b \cup F_{e1} \cup F_{e2}$ and $F^c = F_b$ in Fig. 9(a) and (b), respectively. From Fig. 9, similar to that in Fig. 7, when the colluded copy has high resolution, the selfish colluder can significantly reduce his or her own probability of being detected by reducing the frame rate before multiuser collusion; while when the colluded copy has low resolution, it cannot further lower the selfish colluder's risk. In general, reducing the temporal resolution before collusion can further reduce the selfish colluder's risk only when $F^c \supset \tilde{F}^{(i_1)}$.

D. Performance Comparison of Different Strategies

In the scalable fingerprinting system in Section IV-B, each selfish colluder has two choices when modifying the resolution of his or her fingerprinted copy before collusion. For example, for a selfish colluder $\mathbf{u}^{(i_1 \in \text{SC}^{\text{all}})}$ who receives all three layers from the content owner, during precollusion processing, $\mathbf{u}^{(i_1)}$ can drop the received enhancement layer 2 before collusion and tell other attackers that he or she has a medium-quality fingerprinted copy. $\mathbf{u}^{(i_1)}$ can also drop both enhancement layers and claim that he or she has the base layer only. This section

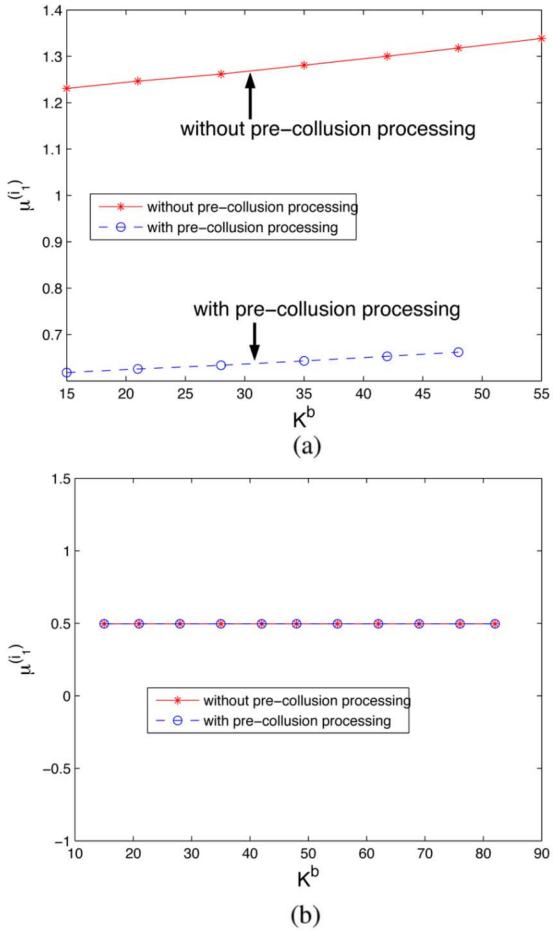

Fig. 9. Comparison of the means of the selfish colluder's detection statistics when he or she reduces the frame rate during precollusion processing with that when he or she does not process his or her copy before multiuser collusion. (a)$F^c = F_b \cup F_{e1} \cup F_{e2}$. (b) $F^c = F_b$. $(|F_b|, |F_{e1}|, |F_{e2}|) = (10, 10, 20)$ and $(N_b, N_{e1}, N_{e2}) = (50000, 50000, 100000)$. There are a total of $M = 450$ users in the system and a total of $K = 150$ colluders. Each K^b represents a unique point on line (17). $\text{CP}^{(i_1)} = (F_b \cup F_{e1} \cup F_{e2}, F_b)$.

compares the effectiveness of different precollusion processing strategies in reducing the selfish colluder's risk, assuming that the quality constraints are satisfied and other colluders do not discover the precollusion processing behavior.

From the analysis in the previous section, neither increasing nor reducing the temporal resolution can further reduce the selfish colluder's probability of being detected when the colluded copy only contains frames in the base layer. Therefore, in this section, we consider scenarios where the colluded copy includes at least one enhancement layer and F^c is equal to either $F_b \cup F_{e1}$ or $F_b \cup F_{e1} \cup F_{e2}$.

Our simulation setup is similar to that in Section IV-C1. We assume that each frame has 5000 embeddable coefficients and we test on a total of 40 frames. We consider a temporally scalable video coding system with $F_b = \{1, 5, 9, \ldots\}$, $F_{e1} = \{3, 7, 11, \ldots\}$ and $F_{e2} = \{2, 4, 8, \ldots\}$, and the lengths of the fingerprints embedded in the base-layer enhancement layer 1 and enhancement layer 2 are $N_b = 50000, N_{e1} = 50000$ and $N_{e2} = 100000$, respectively. We further assume that there are a total of $M = 450$ users in the system, and $|\mathbf{U}^b| = |\mathbf{U}^{b,e1}| = |\mathbf{U}^{\text{all}}| = 150$. For each user, a unique vector

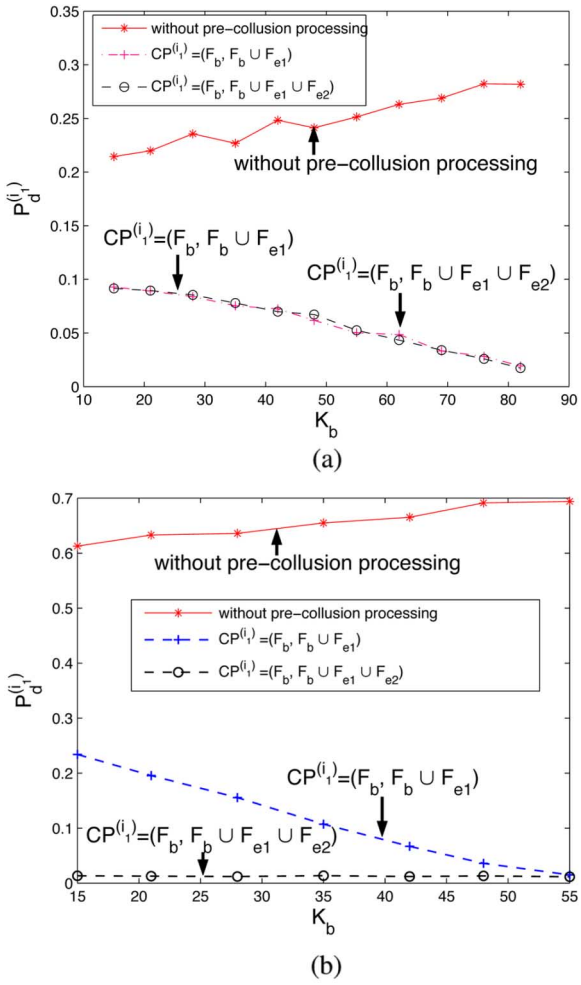

Fig. 10. Performance comparison of different precollusion processing strategies for selfish colluders in SC^b. (a) $F^c = F_b \cup F_{e1}$, (b) $F^c = F_b \cup F_{e1} \cup F_{e2}$. $(|F_b|, |F_{e1}|, |F_{e2}|) = (10, 10, 20)$ and $(N_b, N_{e1}, N_{e2}) = (50000, 50000, 100000)$. Assume that there are a total of $M = 450$ users and $|\mathbf{U}^b| = |\mathbf{U}^{b,e1}| = |\mathbf{U}^{\mathrm{all}}| = 150$. The total number of colluders is $K = 150$ and there is only one selfish colluder. Each value of K^b represents a unique $(K^b, K^{b,e1}, K^{\mathrm{all}})$ on line (17). $P_{fa} = 0.01$ is fixed by selecting the appropriate threshold h in the simulation runs.

is first generated from distribution $\mathcal{N}(0, \sigma_W^2)$ with $\sigma_W^2 = 1/9$, and Gram–Schmidt orthogonalization is applied to let (1) hold strictly for the assigned fingerprints.

During collusion, we assume that there are a total of $K = 150$ colluders and $(K^b, K^{b,e1}, K^{\mathrm{all}})$ takes different values on line (17). We further assume that the additive noise $\mathbf{n}_j$ in (6) follows distribution $\mathcal{N}(0, \sigma_n^2)$ with $\sigma_n^2 = 2\sigma_W^2$. In our simulations, we assume that there is only one selfish colluder $\mathbf{u}^{(i_1)}$ and other colluders do not discover his or her precollusion processing.

1) For Selfish Colluders in Subgroup SC^b*:* For a selfish colluder $\mathbf{u}^{(i_1)}$ who receives the base layer only, $\mathbf{u}^{(i_1)}$ can increase the frame rate of his or her fingerprinted copy with two different parameters $\mathrm{CP}_1^{(i_1)} = (F_b, F_b \cup F_{e1})$ and $\mathrm{CP}_2^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$. In this section, we compare the effectiveness of these two strategies in reducing $\mathbf{u}^{(i_1)}$'s probability of being caught $P_d^{(i_1)}$.

We fix the probability of accusing a given innocent user P_{fa} as 0.01, and compare $P_d^{(i_1)}$ of different precollusion processing parameters. Fig. 10 shows our simulation results when $(K^b, K^{b,e1}, K^{\mathrm{all}})$ takes different values on line (17), and each K^b corresponds to a unique point on that line. $F^c = F_b \cup F_{e1}$ and $F^c = F_b \cup F_{e1} \cup F_{e2}$ in Fig. 10(a) and (b), respectively. From the selfish colluder's point of view, when $F^c = F_b \cup F_{e1}$, the two processing parameters have the same performance. If $F^c = F_b \cup F_{e1} \cup F_{e2}$, $\mathrm{CP}_2^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$ gives the selfish colluder a smaller probability of being detected than $\mathrm{CP}_1^{(i_1)} = (F_b, F_b \cup F_{e1})$. Therefore, under the quality constraints, a selfish colluder in SC^b should pretend to have received all three layers from the content owner in order to minimize his or her risk.

In Fig. 11, we consider two colluders: $\mathbf{u}^{(i_1)}$ who increase the resolution of his or her copy during precollusion processing and $\mathbf{u}^{(i_2)}$ who does not process his or her copy before collusion, and compare their probabilities of being detected by the fingerprint detector. $i_1 \in \mathrm{SC}_b$ and $\mathrm{CP}^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$. From Fig. 11, precollusion processing makes $\mathbf{u}^{(i_2)}$ have a much larger probability of being detected than $\mathbf{u}^{(i_1)}$, and increases $\mathbf{u}^{(i_2)}$'s relative risk when compared with $\mathbf{u}^{(i_1)}$. It is certainly a selfish behavior.

2) For Selfish Colluders in Subgroup $\mathrm{SC}^{b,e1}$*:* For a selfish colluder $\mathbf{u}^{(i_1 \in \mathrm{SC}^{b,e1})}$ who receives the base layer and the enhancement layer 1 from the content owner, $\mathbf{u}^{(i_1)}$ can increase the resolution of his or her copy with parameter $\mathrm{CP}_1^{(i_1)} = (F_b \cup F_{e1}, F_b \cup F_{e1} \cup F_{e2})$ during precollusion processing. $\mathbf{u}^{(i_1)}$ can also drop his or her fingerprinted enhancement layer 1 with parameter $\mathrm{CP}_2^{(i_1)} = (F_b \cup F_{e1}, F_b)$.

From the simulation results shown in Fig. 12(a), when the colluded copy has a medium temporal resolution and $F^c = F_b \cup F_{e1}$, dropping the enhancement layer 1 with parameter $\mathrm{CP}_2^{(i_1)}$ reduces $\mathbf{u}^{(i_1)}$'s probability of being detected, while increasing the resolution with parameter $\mathrm{CP}_1^{(i_1)}$ cannot further lower the selfish colluder's risk. From Fig. 12(b), when the colluded copy includes all three layers and $F^c = F_b \cup F_{e1} \cup F_{e2}$, both $\mathrm{CP}_1^{(i_1)}$ and $\mathrm{CP}_2^{(i_1)}$ can reduce $\mathbf{u}^{(i_1)}$'s probability of being captured, while $\mathrm{CP}_1^{(i_1)}$ gives $\mathbf{u}^{(i_1)}$ a smaller chance to be detected than $\mathrm{CP}_2^{(i_1)}$.

Consequently, in order for a selfish colluder in subgroup $\mathrm{SC}^{b,e1}$ to minimize his or her own risk, when the colluders plan to generate a colluded copy of medium temporal resolution, the selfish colluder should drop the enhancement layer 1 before multiuser collusion; and when the colluders plan to generate a colluded copy containing all three layers, the selfish colluder should increase the resolution of his or her fingerprinted copy with parameter $\mathrm{CP}_1^{(i_1)} = (F_b \cup F_{e1}, F_b \cup F_{e1} \cup F_{e2})$.

Fig. 13 investigates the impact of precollusion processing on other colluders' probability of being detected. In Fig. 13, there are ten selfish colluders who use the same parameter $(F_b \cup F_{e1}, F_b \cup F_{e1} \cup F_{e2})$ during precollusion processing, and they process their fingerprinted copies independently. We consider two colluders—a selfish colluder $\mathbf{u}^{(i_1)}$ and another colluder $\mathbf{u}^{(i_2)}$ who does not apply precollusion processing. In this scenario, precollusion processing not only reduces the selfish colluders' absolute risk, it also decreases other attackers'

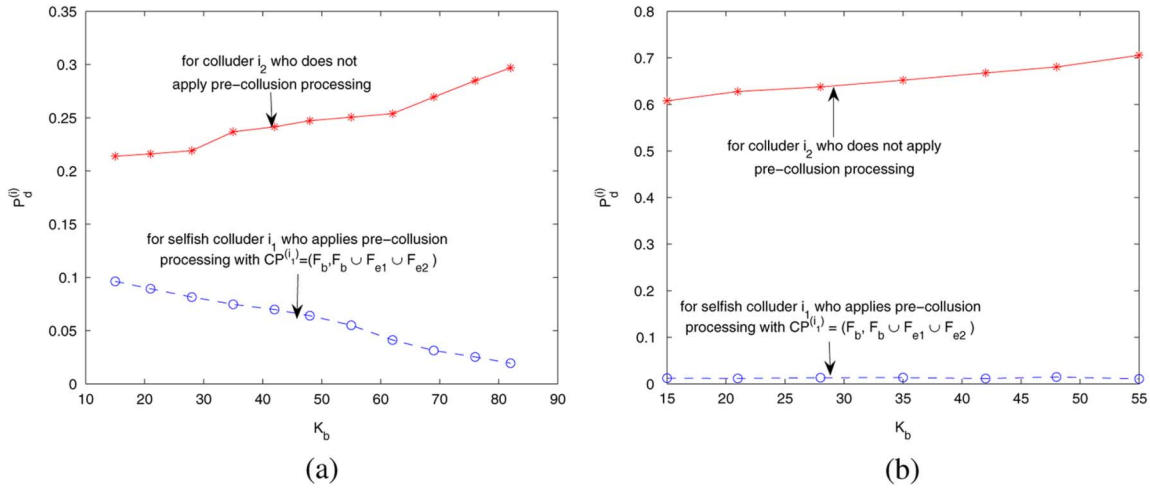

Fig. 11. Comparison of different colluders' probabilities of being detected when selfish colluders exist. (a) $F^c = F_b \cup F_{e1}$. (b) $F^c = F_b \cup F_{e1} \cup F_{e2}$. $\mathbf{u}^{(i_1)}$ is a selfish colluder, and $\mathbf{u}^{(i_2)}$ does not process his or her copy before multiuser collusion. $\mathrm{CP}^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$. The simulation setup is the same as in Fig. 10.

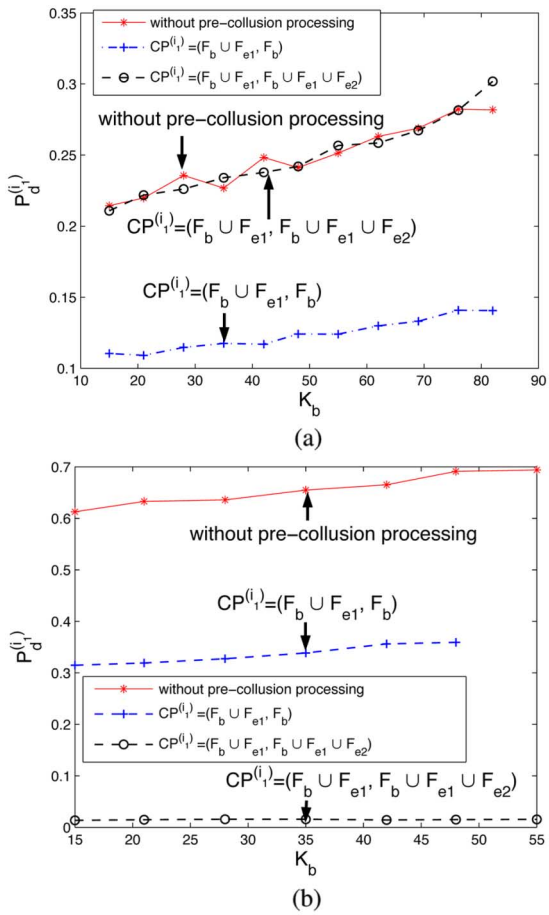

Fig. 12. Performance comparison of different precollusion processing strategies for selfish colluders in $\mathrm{SC}^{b,e1}$. (a) $F^c = F_b \cup F_{e1}$. (b) $F^c = F_b \cup F_{e1} \cup F_{e2}$. $(|F_b|, |F_{e1}|, |F_{e2}|) = (10, 10, 20)$ and $(N_b, N_{e1}, N_{e2}) = (50000, 50000, 100000)$. Assume that there are a total of $M = 450$ users and $|\mathbf{U}^b| = |\mathbf{U}^{b,e1}| = |\mathbf{U}^{\mathrm{all}}| = 150$ $K = 150$, and assume that there is only one selfish colluder. Each K^b corresponds to a unique triplet $(K^b, K^{b,e1}, K^{\mathrm{all}})$ on line (17). We select the threshold to fix $P_{fa} = 0.01$.

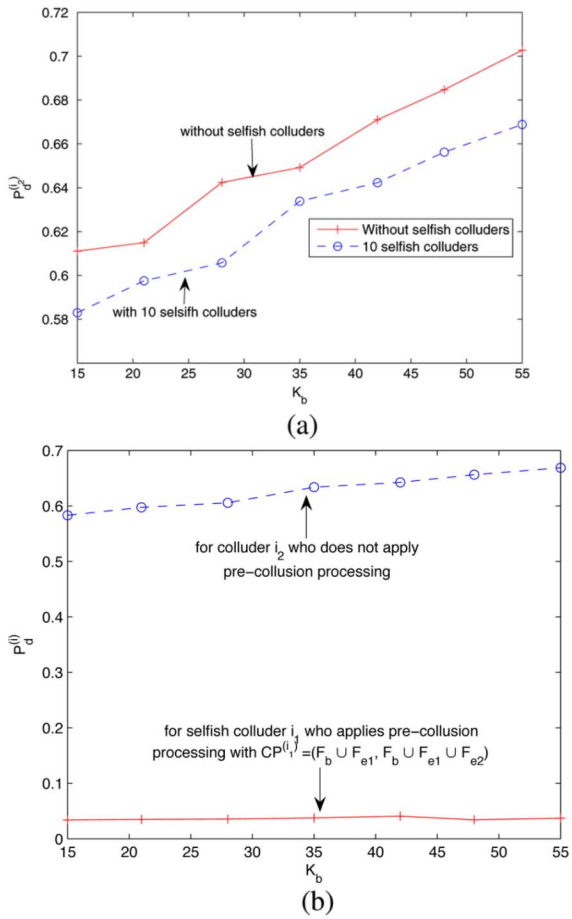

Fig. 13. Impact of precollusion processing on other colluders' probability of being detected. $\mathbf{u}^{(i_1)}$ is a selfish colluder and colluder $\mathbf{u}^{(i_2)}$ does not process his or her copy before collusion. (a) $\mathbf{u}^{(i_2)}$'s probability of being detected $(P_d^{(i_2)})$. (b) Comparison of $\mathbf{u}^{(i_1)}$'s probability of being detected with that of $\mathbf{u}^{(i_2)}$. The simulation setup is the same as that in Fig. 12(b). There are ten selfish colluders who select the same parameter $(F_b \cup F_{e1}, F_b \cup F_{e1} \cup F_{e2})$ during precollusion processing, while each processes his or her copy independently. The colluded copy has the highest temporal resolution with $F^c = F_b \cup F_{e1} \cup F_{e2}$.

probability of being detected. However, from Fig. 13(b), such precollusion processing makes $\mathbf{u}^{(i_1)}$ take a much smaller chance of being caught than $\mathbf{u}^{(i_2)}$ and increases other colluders' relative risk with respect to the selfish colluders. Therefore, it is still a selfish behavior.

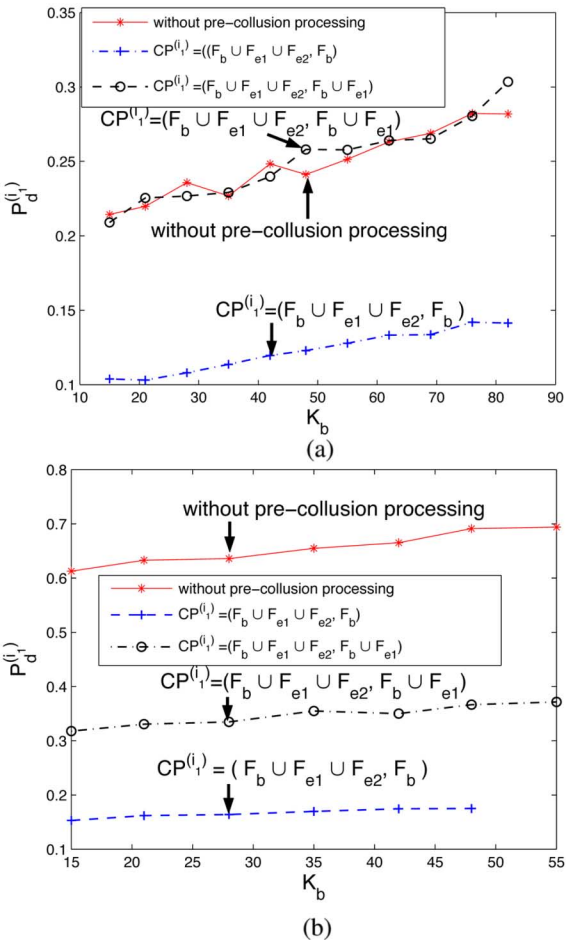

Fig. 14. Performance comparison of different precollusion processing strategies for selfish colluders in $\mathrm{SC}^{\mathrm{all}}$. (a) $F^c = F_b \cup F_{e1}$. (b)$F^c = F_b \cup F_{e1} \cup F_{e2}$. $(|F_b|, |F_{e1}|, |F_{e2}|) = (10, 10, 20)$ and $(N_b, N_{e1}, N_{e2}) = (50000, 50000, 100000)$. $M = 450$ and $|\mathbf{U}^b| = |\mathbf{U}^{b,e1}| = |\mathbf{U}^{\mathrm{all}}| = 150$. The total number of colluders is $K = 150$ and assume that there is only one selfish colluder. Each K^b corresponds to a unique triplet $(K^b, K^{b,e1}, K^{\mathrm{all}})$ on line (17). The threshold h is selected to fix $P_{fa} = 0.01$.

3) For Selfish Colluders in $\mathrm{SC}^{\mathrm{all}}$*:* For a selfish colluder $\mathbf{u}^{(i_1)}$ in subgroup $\mathrm{SC}^{\mathrm{all}}$ who receives all three layers, during precollusion processing, $\mathbf{u}^{(i_1)}$ can reduce the frame rate of his or her fingerprinted copy with two different parameters $\mathrm{CP}_1^{(i_1)} = (F_b \cup F_{e1} \cup F_{e2}, F_b)$ and $\mathrm{CP}_2^{(i_1)} = (F_b \cup F_{e1} \cup F_{e2}, F_b \cup F_{e1})$.

As shown in Fig. 14(a), when the colluded copy has medium resolution, using $\mathrm{CP}_1^{(i_1)}$ further reduces $\mathbf{u}^{(i_1)}$'s probability of being detected, while $\mathrm{CP}_2^{(i_1)}$ does not change his or her risk. From Fig. 14(b), if the colluders generate a high-resolution colluded copy, both strategies lower the selfish colluder's probability of being captured and $P_d^{(i_1)}$ of $\mathrm{CP}_1^{(i_1)}$ is smaller than $P_d^{(i_1)}$ of $\mathrm{CP}_2^{(i_1)}$. Consequently, from the selfish colluder's point of view, dropping both enhancement layers before multiuser collusion is preferred for a selfish colluder in subgroup $\mathrm{SC}^{\mathrm{all}}$ to minimize his or her risk of being detected.

For colluder $\mathbf{u}^{(i_2)}$, who does not process his or her received copy before collusion, Fig. 15 shows the impact of the selfish colluders' precollusion processing on $\mathbf{u}^{(i_2)}$'s probability of being detected when the total number of selfish colluders varies.

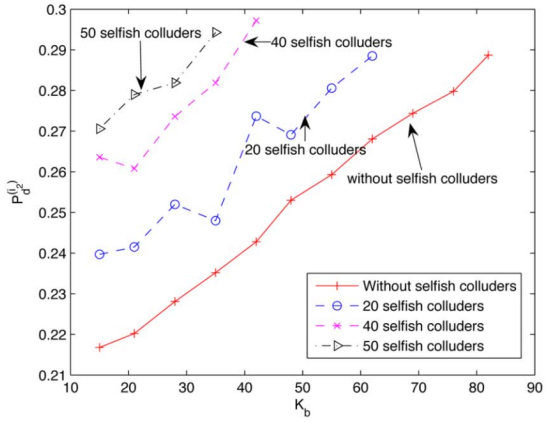

Fig. 15. For colluder $\mathbf{u}^{(i_2)}$, who does not apply precollusion processing, $\mathbf{u}^{(i_2)}$'s probability of being detected has a different number of selfish colluders. The simulation setup is the same as in Fig. 14(a) and $F^c = F_b \cup F_{e1}$. All selfish colluders select the same parameter $\mathrm{CP} = (F_b \cup F_{e1} \cup F_{e2}, F_b)$, while each processes his or her own copy independently.

In Fig. 15, all selfish colluders select the same precollusion processing parameter $\mathrm{CP} = (F_b \cup F_{e1} \cup F_{e2}, F_b)$, while each processes his or her fingerprinted copy independently. From Fig. 15, dropping enhancement layers before collusion increases others' probability of being detected, and $\mathbf{u}^{(i_2)}$ has a larger probability of being detected when there are more selfish colluders. In this example, precollusion processing is not only selfish, but also malicious.

E. Simulation Results on Real Video

We test the effectiveness of changing the resolution of the fingerprinted copy before collusion on real videos, assuming that the quality constraints are satisfied. We choose the first 40 frames of sequence "car phone" as an example. Similar to that in Section IV-C1, we consider a temporally scalable video coding system with $F_b = \{1, 5, 9, \ldots\}$, $F_{e1} = \{3, 7, 11, \ldots\}$ and $F_{e2} = \{2, 4, 8, \ldots\}$. The lengths of the fingerprints embedded in the base layer, enhancement layer 1, and enhancement layer 2 are $N_b = 39988, N_{e1} = 39934$, and $N_{e2} = 79686$, respectively. We assume that the total number of users is $M = 450$ and $|\mathbf{U}^b| = |\mathbf{U}^{b,e1}| = |\mathbf{U}^{\mathrm{all}}| = 150$. We adopt the human visual model-based spread-spectrum embedding in [17], and embed the fingerprints in the DCT domain. We first generate independent vectors following distribution $\mathcal{N}(0, 1/9)$, and then apply Gram–Schmidt orthogonalization to let the assigned fingerprints be strictly orthogonal and have equal energy. In each fingerprinted copy, similar to that in [20], fingerprints in adjacent frames are correlated with each other, depending on the similarity between the host frames.

During collusion, we assume that there are a total of $K = 150$ colluders, and $(K^b, K^{b,e1}, K^{\mathrm{all}})$ takes different values on line (17). We consider a simple scenario where there is only one selfish colluder who changes the resolution of his or her received copy before collusion. Furthermore, we assume that no colluders discover the selfish colluder's precollusion processing. In our simulations, we adjust the power of the additive noise $\mathbf{n}_j$ such that $\|\mathbf{n}_j\|^2 / \|\mathrm{JND}_j \cdot \mathbf{W}_j^{(i)}\|^2 = 2$ for every frame in the colluded copy.

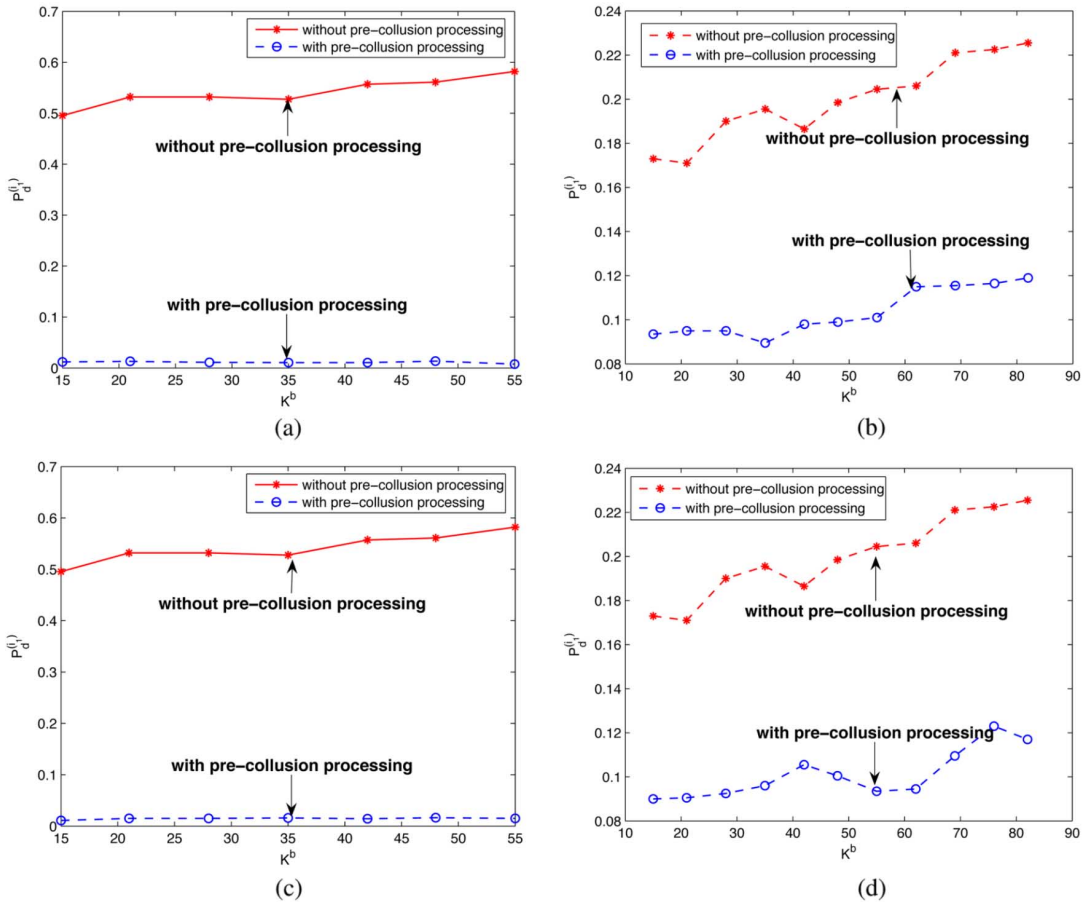

Fig. 16. Simulation results of changing the resolution of the received copies during precollusion processing on the first 40 frames of sequence car phone. (a) $CP^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$. (b): $CP^{(i_1)} = (F_b \cup F_{e1}, F_b)$. (c): $CP^{(i_1)} = (F_b \cup F_{e1}, F_b \cup F_{e1} \cup F_{e2})$. (d): $CP^{(i_1)} = (F_b \cup F_{e1} \cup F_{e2}, F_b)$ $(|F_b|, |F_{e1}|, |F_{e2}|) = (10, 10, 20)$. The total number of users is $M = 450$ and $|\mathbf{U}^b| = |\mathbf{U}^{b,e1}| = |\mathbf{U}^{\text{all}}| = 150$. There are a total of $K = 150$ colluders and each K^b represents a unique point on line (17). P_{fa} is fixed as 10^{-2} by selecting the threshold h. In (a) and (c), $F^c = F_b \cup F_{e1} \cup F_{e2}$. In (b) and (d), $F^c = F_b \cup F_{e1}$.

Fig. 16 shows the simulation results. In Fig. 16(a), the selfish colluder $\mathbf{u}^{(i_1)}$ receives the fingerprinted base layer only and he or she increases the frame rate with parameter $\text{CP}^{(i_1)} = (F_b, F_b \cup F_{e1} \cup F_{e2})$ before multiuser collusion. In Fig. 16(b) and (c), the selfish colluder receives a copy of medium resolution from the content owner and the precollusion processing parameters are $\text{CP}^{(i_1)} = (F_b \cup F_{e1}, F_b)$ and $\text{CP}^{(i_1)} = (F_b \cup F_{e1}, F_b \cup F_{e1} \cup F_{e2})$, respectively. In Fig. 16(d), the selfish colluder receives from the content owner all three layers, and he or she drops both enhancement layers during precollusion processing. In Fig. 16(a) and (c), $F^c = F_b \cup F_{e1} \cup F_{e2}$ and the colluded copy has the highest frame rate; and the colluded copy has medium temporal resolution and $F^c = F_b \cup F_{e1}$ in Fig. 16(b) and (d).

From Fig. 16, under the quality constraints, changing the resolution of the fingerprinted copy can help a selfish colluder further reduce his or her risk of being caught, especially when the colluded copy has high resolution. The simulation results on real videos agree with our theoretical analysis in Section IV-C, and are comparable with the results in Section IV-D.

V. Conclusions and Discussion

In this paper, we consider the problem of traitors within traitors in behavior forensics and formulate the dynamics among attackers during collusion to minimize their own risk of being detected and protect their own interest. As the first work on the analysis of the behavior dynamics during collusion, we investigate a few precollusion processing strategies that a selfish colluder can use to further reduce his or her chance of being captured by the digital rights enforcer, and analyze their effectiveness. We also analyze the constraints on pre-collusion processing to maintain the perceptual quality of the fingerprinted copies.

We first investigate the strategies for a selfish colluder to attenuate the energy of the embedded fingerprints before collusion. The selfish colluder can apply temporal filtering to his or her copy and average adjacent frames of similar content before multiuser collusion. We analyze its effectiveness in reducing the selfish colluder's risk as well as the perceptual quality of the fingerprinted copy after temporal filtering. Both our analytical and simulation results show that this temporal filtering reduces the selfish colluder's risk of being captured at the cost of quality degradation. We then investigate the tradeoff between the risk and the perceptual quality that a selfish colluder needs to address, and derive the optimal filtering coefficients to minimize his or her probability of being caught while maintaining good quality of his or her fingerprinted copy.

We then consider the problem of traitors within traitors when attackers receive fingerprinted copies of different resolutions due to network and device heterogeneity. In such a scenario, in

addition to temporal filtering, a selfish colluder can also change the resolution and quality of his or her fingerprinted copy before multiuser collusion. We show that under the quality constraints, changing the resolution of the fingerprinted copy can help a selfish colluder further reduce his or her probability of being caught, especially when the colluded copy has high quality. For traitors within traitors in scalable fingerprinting systems, we also investigate the selection of the optimal strategy for a selfish colluder to minimize his or her risk under the quality constraints.

The work described in this paper is an initial step toward a thorough investigation of traitor-within-traitor behavior forensics. This paper focuses on one aspect of precollusion processing (i.e., the risk that colluders are detected by the content owner). Future work might address the risk that selfish traitors are detected by their fellow traitors, because the statistics of their preprocessed copies differ from those of unprocessed ones. The natural framework for such a study is a game theoretic one, in which admissible strategies and cost functions are defined, and the strategy that maximizes the payoff function for the coalition is derived. Among other benefits, such a framework would inherently address the issue of detecting selfish behavior by the fellow colluders and seek whether there exists an equilibrium from which no colluder has an interest to deviate—in this case, selfish behavior would be counterproductive.

REFERENCES

[1] I. Cox, J. Killian, F. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.

[2] H. Stone, Analysis of Attacks on Image Watermarks With Randomized Coefficients NEC Res. Inst., 1996, Tech. rep. 96-045.

[3] I. Cox and J. P. Linnartz, "Some general methods for tampering with watermarking," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 4, pp. 587–593, May 1998.

[4] F. Petitcolas, R. Anderson, and M. Kuhn, "Attacks on copyright marking systems," in *Proc. 2nd Workshop Info. Hiding, Lecture Notes Comput. Sci.*, Apr. 1998, pp. 218–238.

[5] F. Hartung, J. Su, and B. Girod, "Spread spectrum watermarking: Malicious attacks and counterattacks," in *Proc. SPIE, Security and Watermarking of Multimedia Contents, Electronic Imaging*, Jan. 1999, pp. 147–158.

[6] F. Zane, "Efficient watermark detection and collusion security," in *Proc. 4th Int. Conf. Financial Cryptography, Lecture Notes Comput. Sci.*, Feb. 2000, vol. 1962, pp. 21–32.

[7] J. Dittmann, P. Schmitt, E. Saar, J. Schwenk, and J. Ueberberg, "Combining digital watermarks and collusion secure fingerprints for digital images," *SPIE J. Electron. Imaging*, vol. 9, no. 4, pp. 456–467, Oct. 2000.

[8] W. Trappe, M. Wu, Z. Wang, and K. J. R. Liu, "Anti-collusion fingerprinting for multimedia," *IEEE Trans. Signal Processing*, vol. 51, no. 4, pp. 1069–1087, Apr. 2003.

[9] Z. J. Wang, M. Wu, W. Trappe, and K. J. R. Liu, "Group-oriented fingerprinting for multimedia forensics," *EURASIP J. Appl. Signal Processing*, vol. 2004, no. 14, pp. 2142–2162, Nov. 2004.

[10] F. Ergun, J. Killian, and R. Kumar, "A note on the limits of collusion-resistant watermarks," *Advances in Cryptology, Lect. Notes Comput. Sci.*, vol. 1592, pp. 140–149, 2001.

[11] J. Killian, T. Leighton, L. R. Matheson, T. G. Shamoon, R. Tajan, and F. Zane, Resistance of Digital Watermarks to Collusive Attacks Dept. Comput. Sci., Princeton Univ., Princeton, NJ, 1998, Tech. Rep. TR-585-98.

[12] J. Su, J. Eggers, and B. Girod, "Capacity of digital watermarks subject to an optimal collusion attacks," presented at the Eur. Signal Processing Conf., Tampere, Finland, 2000.

[13] H. Zhao, M. Wu, Z. J. Wang, and K. J. R. Liu, "Forensic analysis of nonlinear collusion attacks for multimedia fingerprinting," *IEEE Trans. Image Process.*, vol. 14, no. 5, pp. 646–661, May 2005.

[14] D. Schonberg and D. Kirovski, "Fingerprinting and forensic analysis of multimedia," in *Proc. Multimedia 04: Proc. 12th Annu. ACM Int. Conf. Multimedia*, Oct. 2004, pp. 788–795.

[15] Z. J. Wang, M. Wu, H. Zhao, W. Trappe, and K. J. R. Liu, "Anti-collusion forensics of multimedia fingerprinting using orthogonal modulation," *IEEE Trans. Image Process.*, vol. 14, no. 6, pp. 804–821, Jun. 2005.

[16] H. Zhao and K. J. R. Liu, "Behavior forensics for scalable multi-user collusion: Fairness and effectiveness," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 3, pp. 311–329, Sep. 2006.

[17] C. Podilchuk and W. Zeng, "Image adaptive watermarking using visual models," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 4, pp. 525–540, May 1998.

[18] M. Swanson, B. Zhu, and A. Tewfik, "Multiresolution scene-based video watermarking using perceptual models," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 4, pp. 540–550, May 1998.

[19] K. Su, D. Kundur, and D. Hatzinakos, "Statistical invisibility for collusion-resistant digital video watermarking," *IEEE Trans. Multimedia*, vol. 7, no. 1, pp. 43–51, Feb. 2005.

[20] K. Su, D. Kundur, and D. Hatzinakos, "Spatially localized image-dependent watermarking for statistical invisibility and collusion resistance," *IEEE Trans. Multimedia*, vol. 7, no. 1, pp. 52–66, Feb. 2005.

[21] S. Yoon and N. Ahuja, "Frame interpolation using transmitted block-based motion vectors," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2001, vol. 3, pp. 856–859.

[22] Y. Wang, J. Ostermann, and Y. Zhang, *Video Processing and Communications*, 1st ed. Upper Saddle River, NJ: Prentice-Hall, 2001.

H. Vicky Zhao (M'05) received the B.S. and M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1997 and 1999, respectively, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 2004.

She has been a Research Associate with the Department of Electrical and Computer Engineering and the Institute for Systems Research, University of Maryland. Since 2006, she has been an Assistant Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. She coauthored the book *Multimedia Fingerprinting Forensics for Traitor Tracing* (Hindawi, 2005). Her research interests include information security and forensics, multimedia, digital communications, and signal processing.

K. J. Ray Liu (F'03) is Professor and Associate Chair, Graduate Studies and Research of the Electrical and Computer Engineering Department, University of Maryland, College Park. His research contributions encompass broad aspects of wireless communications and networking, information forensics and security, multimedia communications and signal processing, bioinformatics and biomedical imaging, and signal processing algorithms and architectures.

Dr. Liu is the recipient of best paper awards from the IEEE Signal Processing Society (twice), IEEE Vehicular Technology Society, and EURASIP, IEEE Signal Processing Society Distinguished Lecturer, EURASIP Meritorious Service Award, and the National Science Foundation Young Investigator Award. He also received the Poole and Kent Company Senior Faculty Teaching Award and Invention of the Year Award, both from the University of Maryland. He is Vice President—Publications and on the Board of Governor of the IEEE Signal Processing Society. He was the Editor-in-Chief of *IEEE Signal Processing Magazine* and the founding Editor-in-Chief of *EURASIP Journal on Applied Signal Processing*.