Applied Soft Computing 48 (2016) 240–253

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

On digital image trustworthiness

Donghui Hu[a,*], Xiaotian Zhang[a], Yuqi Fan[a], Zhong-Qiu Zhao[a], Lina Wang[b], Xintao Wu[c], Xindong Wu[d]

[a] College of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230009, China
[b] College of Computer Science, Wuhan University, Wuhan 430072, China
[c] Department of Computer Science and Computer Engineering, University of Arkansas, Fayetteville, AR 72701, USA
[d] Department of Computer Science, University of Vermont, Burlington, VT 05405, USA

ARTICLE INFO

Article history:
Received 6 October 2014
Received in revised form 10 January 2016
Accepted 2 July 2016
Available online 16 July 2016

Keywords:
Digital image
Trustworthiness evaluation
Dempster-Shafer theory
Information fusion

ABSTRACT

Digital images are facing a crisis of trustworthiness with the emergence of various digital image processing and steganography tools. This paper proposes a novel approach that can evaluate the trustworthiness of a digital image. In this approach, an information fusion method is used to combine base digital image forensic models at the feature level and the decision level. When using different kinds of base forensic models to get supporting evidence for different kinds of digital image manipulations, there exist uncertainties introduced by base forensic models and conflicts among evidence provided by different forensic models. We use the Dempster-Shafer (D-S) evidence theory and an improved least square method to tolerate the uncertainties of forensic models and reduce the evidence conflicts. The lower and upper limits of digital image trustworthiness can then be reliably evaluated by the D-S theory. Three information fusion models based on the D-S theory are proposed. The first model uses the D-S theory at the feature fusion level. The second uses the D-S theory at the decision fusion level, where an improved least square method is designed to reduce the evidence conflicts. The last model is a combination of the first and the second one, where the D-S theory is applied at both the feature fusion and decision fusion levels. Experiments are carried out on four kinds of digital image manipulations. The experimental results show that the three proposed models are very stable in evaluating different kinds of natural images and tampering images. While the first model can only give the upper limit of the trustworthiness of a digital image, the second and the third one can give both lower and upper limits of the trustworthiness of a digital image, as well as the uncertainties of the evidence produced by base forensics models. Compared with the second model, the third one can further reduce the uncertainties. The experimental results also show that when a digital image undergoes many kinds of manipulations, our models can validly compute a soft degree to measure the trustworthiness of the image, while current ordinary digital image forensic models may fail to predict it correctly. Experimental results also demonstrate that the proposed digital image trustworthiness evaluation models can be adapted as digital image forensic classification models with very high detection accuracy.

1. Introduction

As the representation of natural scenes, digital images traditionally imply the truth and fact. However, with the popularization of various digital image processing tools, people can easily manipulate digital images. Digital images can also be used to hide secret information by various steganography tools. Those manipulations greatly undermine the credibility of digital images. Nowadays, people are often concerned about whether a digital image is trustworthy before they are comfortable to use it. As a result, digital images are facing crisis of trustworthiness.

Trust is a common phenomenon of our human life. It has been argued that we as humans would not even be able to face the complexities of the world without resorting to trust [1]. Many literatures have modeled trust from different perspectives. The Oxford Reference Dictionary states the trust as the firm belief in the reliability or truth or strength of an entity. Golembiewski et al. [2] think that trust "implies some degree of uncertainty as to outcome", and trust also "implies hopefulness or optimism as to outcome".

* Corresponding author.
E-mail addresses: hudh@hfut.edu.cn (D. Hu), msadsl6233925@126.com (X. Zhang), yuqi.fan@gmail.com (Y. Fan), zhongqiuzhao@gmail.com (Z.-Q. Zhao), lnawang@163.com (L. Wang), xintaowu@uark.edu (X. Wu), xwu@cs.uvm.edu (X. Wu).

http://dx.doi.org/10.1016/j.asoc.2016.07.010
1568-4946/

Yamamoto et al. [3] state that "the decision to trust is based on evidence to believe, or be confident in, someone's or something's good intentions towards us". Grandison et al. [4] consider that trust is a composition of many different attributes – reliability, dependability, honesty, truthfulness, security, competence, and timeliness, which may have to be considered depending on the environment in which trust is being specified.

In this paper, we consider the scenario that, when people or software use a digital image, they may be concerned about its integrity, origin and/or security. People tend to use more trustworthy images than un-trustworthy ones. Inspired by the trust definition of Grandison et al. [4], in this paper, we propose models to measure the trustworthiness of a digital image using different trustworthiness attributes. We term our models as digital image trustworthiness evaluation models. The trustworthiness of a digital image is a value that tells how trustworthy the image can be used. Digital images may have many trustworthiness attributes or trustworthiness evaluation indexes, such as the index of security (to evaluate whether the image contains hidden secret information), the index of integrity (to evaluate whether the digital image is manipulated), the index of authenticity (to evaluate whether the image is generated by a digital camera or by a computer graphics render tool), and so on. Different digital image users may have concerns with different trustworthiness attributes. In this paper, we propose models to effectively calculate the degree of a digital image's trustworthiness.

Many techniques associated with trustworthiness of digital images have been developed in recent years. Active techniques such as digital watermarking [5] and perceptual image hash [6,7] can ensure the trustworthiness of a digital image. In these cases, the original image must be in hand, and the corresponding digital watermarking should be embedded (or perceptual image hash should be generated). However, in real life, most of the digital images do not contain digital watermarking nor have an image hash. Digital image forensics [8] is a passive technique which can determine whether an image is tampered or not. However, it cannot capture the degree of trustworthiness for a given image, nor can measure trustworthy attributes (e.g. the origin, integrity and security) synthetically. In this paper, an effective method that can measure the trustworthiness of a digital image synthetically and quantitatively will be developed.

Moreover, in this paper, when designing methods to evaluate the trustworthiness of a digital image, we use various digital image forensic models to capture the image's different trustworthiness attributes. We find that the evidence supported by different forensic models may conflict with each other, and each forensic model comes with uncertainties when predicting multi-manipulations [9]. To achieve more reliable and effective digital image trustworthiness evaluation, we propose the use of the Dempster-Shafer (D-S) evidence theory and data fusion. The D-S theory [10,11], which is generally used to combine separate pieces of evidence to calculate the probability of an event, is a mathematical theory of evidence based on belief functions and plausible reasoning. The D-S evidence theory allows the representation of ignorance by giving support to the case of "unknown". It can reduce the uncertainty by allowing evidence to support several mutually exclusive conclusions, and can represent data imperfections without the need to make simplifying assumptions [12]. Data fusion is a method to combine data from multiple sensors in order to achieve better accuracy and more specific inferences than those using a single sensor alone [13]. It can be applied at the raw data level, feature level, or decision level [13,14].

Three fusion models are presented based on the D-S theory and data fusion. The first uses feature fusion to address the problem of evidence conflict when using Dempster's rule of combination. The second uses the Dempster's rule of combination at the decision fusion level, in which we develop an improved least square method to deal with the evidence matrix and reduce the conflicts among evidence provided by different forensic models. Uncertainty of each forensic model is calculated in the training phase and then combined in the evaluation phase. The lower and upper limits of digital image trustworthiness are calculated by using the Dempster's rule of combination. The last model is a combination of the first and second ones where the uncertainty is further reduced by the feature fusion.

Finally, we conduct simulations on 11,200 test images. Simulation results show that the proposed three models can evaluate digital image trustworthiness effectively and reliably, and the last one performs the best.

The rest of this paper is organized as follows. Section 2 reviews related work and compares the trustworthiness evaluation methods with current digital forensic ones. Section 3 describes three models based on the D-S theory and information fusion. The effectiveness of the three models is evaluated and compared in Section 4. Finally, conclusions are drawn in Section 5.

2. Related work

Digital image trustworthiness evaluation models are interactively related to digital image forensic models, as shown in Fig. 1. In our proposed digital image trustworthiness evaluation models, we use digital image forensic models to capture different aspects of tampering features and provide forensic decision or evidence for each kind of tampering. The digital image trustworthiness

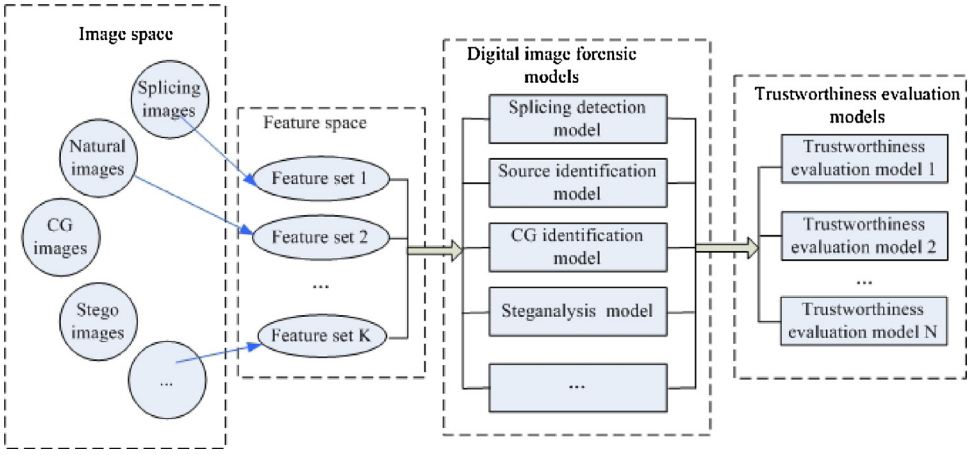

Fig. 1. Relationship between digital image trustworthiness evaluation models and digital image forensic models.