A two-step watermarking attack using long-range correlation image restoration

Amir H. Taherinia* and Mansour Jamzad
Department of Computer Engineering, Sharif University of Technology, Azadi Ave., Tehran, Iran

ABSTRACT
This paper presents an efficient scheme for blind watermark attacking using the concept of matching of the long-range data. The main idea of the proposed attack is to add plenty of noise to the watermarked image and then try to restore an unwatermarked copy of the noisy image. The aim is to destroy the watermark information without accessing the parameters used during the watermark embedding process. So, it allows our approach to be completely free from any pre-assumption on the watermarking algorithm or any other parameters that is used during the watermark embedding procedure. Experimental results show the proposed algorithm’s superiority over several other traditional watermarking benchmarks such as Stirmark and Optimark. Peak signal-to-noise ratio of the watermarked image after applying the proposed attack is more than 45 dB, and the normalized cross-correlation for the extracted watermark is lower than 0·4, so the watermark is not detectable after our attack. Copyright © 2011 John Wiley & Sons, Ltd.

KEYWORDS
image restoration; noise addition watermarking; watermarking attacks; benchmarks

*Correspondence
Amir Taherinia, Department of Computer Engineering, Sharif University of Technology, Azadi Ave., Tehran, Iran.
E-mail: taherinia@ce.sharif.edu

1. INTRODUCTION
Nowadays, in the age of digitally computerized world, digital media is replacing the traditional media and becoming more popular in our daily lives. Computers with Internet connections have made the distribution of digital products much easier. Meantime, we have witnessed a rise in copyright encroachment. Digital watermarking has emerged as a possible solution for authorship proofs or dispute resolving and has become an area of increased research activity. In these applications, there are several requirements that watermarking schemes must fulfill, such as imperceptibility, robustness to attacks that try to destroy a legally inserted watermark or to embed an illegal watermark in some asset.

For verifying the security and robustness of watermarking algorithms, specific attacks have to be applied to test them. A successful attack should damage or destroy the watermark while preserving the commercial quality of the image. Therefore, the attacker should introduce distortions that are limited.

Digital watermark benchmarking refers to the process of impairing the detection of hidden watermark from the digital object (or the host object). Existing watermark benchmarking schemes based on the watermark-impairing method used can be classified into two major categories:

(1) blind attacks, in which the watermark benchmark does not exploit the host signal information during watermark-impairing process;

(2) informed attacks, in which the watermark benchmark exploits knowledge of the watermarked signal or watermarking method during watermark-impairing process.

Early benchmarking tools reported in the literature for evaluating watermarking techniques, such as those of Stirmark [1], Checkmark [2], and Certimark [3], fall into the first category and principally relied on a number of common image processing operations such as geometrical transformations and signal manipulations. Unfortunately, these benchmarks usually degrade the quality of processed image so that they do not permit further commercial exploitation of the attacked image. The latest arrival in the benchmarking field is Optimark [4], which provides an improved graphical interface, but it implements the same attacks as Stirmark, in addition to the possibility of creating combinations of them. The main drawback of the existing benchmarks is not taking into account the statistical