Chapter 3
Online Search and Verification of Watermarked Images for Copyright Infringement Detection

3.1 Overview of Proposed Method

In Section 3.1.1, the motivation of the proposed method for online search and verification of watermarked images for copyright infringement detection is described. And in Section 3.1.2, we briefly describe the principle of the proposed method. In Section 3.2, the proposed method for online search and verification of watermarked images is described. Finally, in Section 3.3, some discussions and a summary are made.

3.1.1 Motivation

With the rapid growth of digital world, how to protect the copyright of digital multimedia has become an urgent issue. And many researches on digital watermarking have been proposed for copyright protection in recent years. It is a pity that we have many novel techniques but lack a great platform to integrate these techniques to facilitate protection of copyright in people's daily life. So we integrate six adopted methods proposed in [1] and [11], which include visible and invisible watermarking techniques, for copyright protection into our system.

The central image authentication center provides local image authentication

centers with two software packages. One software package, called InfoProtector, has implemented six adopted techniques just mentioned above, and can be used for watermark embedding in BMP, GIF and JPEG image formats. The other software package has implemented proposed methods in this chapter. There will be four main functions in the software package, namely, online search and verification of watermarked images for copyright infringement detection, fast watermark verification by progressive image matching, and lossless visible watermarking.

Local image authentication centers can embed their watermarks into their own images using the software InfoProtector. If they find that someone posts their images on the Internet without permission, they can search and verify the ownership of the images immediately by using the latter software package mentioned above and request the central image authentication center to issue a certificate proving the infringement. It is more efficient than the traditional authentication mechanism in which it is necessary to go to the authentication center in person. And illegal users will have no time to destroy the evidence of their infringement.

3.1.2 Principle of Proposed Method

In the proposed hierarchical centers, there are a central image authentication center and many local image authentication centers. We enable all of these centers to search and verify watermarked images due to the enhancement of flexibility of our system.

Before requesting the central image authentication center to issue a certificate, local image authentication centers can search and verify watermarked images in advance if the network is heavy loaded or the central image authentication center is weak in computing. Or local image authentication centers can just give a URL of a webpage or a public FTP server to the central image authentication center if the

network is slightly loaded or the central image authentication center is strong in computing. And the central image authentication center will do all the task of searching, verifying, authenticating and issuing certificates.

Even a person can request by e-mail the central image authentication center for a simplified software package, which only contains the functions of online search of suspected images and verification of watermarked images. This simplified software package can be used to check the legality of images in order to prevent misuse of copyrighted images.

If we find suspected images in a webpage or in a public FTP server when browsing on the Internet, we can download all images in the webpage by parsing the web page source code or download all images in the public FTP server by parsing the directories. And then we verify the owners of these images to see if the copyright of these images have been infringed.

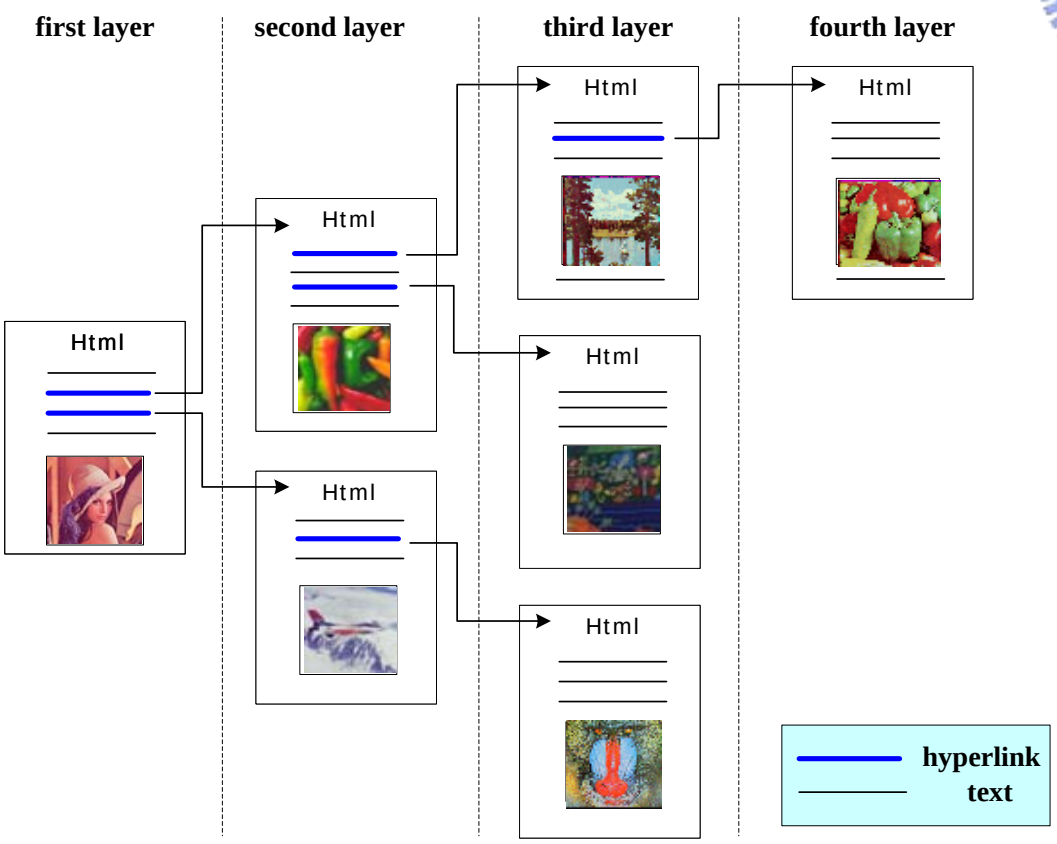
3.2 Proposed Method for Search and Verification of Watermarked Images

In Section 3.2.1, the employed method for detecting images in webpages is introduced. And we describe the proposed method for detecting images in public FTP servers in Section 3.2.2. The process for verification of watermarked images is presented in Section 3.2.3. Finally, the overall procedure for proposed online search and verification method is described in Section 3.2.4.

3.2.1 Employed Method for Detecting Images in Webpages

To find suspected images in a webpage, we can search all images in the webpage by parsing the webpage source code. In other words, we go through every HyperText Markup Language (HTML) tag of the webpage and download an image when we find an html tag indicating there is an image.

If there is an html tag indicating a hyperlink, we have two different methods to handle with. We ignore the hyperlink in a *passive mode* while we deal with the webpage which is linked from the hyperlink further in an *active mode*. That is, we download only images in the webpage of the first layer in the passive mode. But in the active mode, we download all images in not only the webpage of the first layer but also the webpages to which are linked from references in the webpage of the first layer. We download all images in webpages in a depth-first search order until the specified search layer is reached. An example of detecting images in webpages within the fourth layer using the depth-first search algorithm is shown in Figure 3.1(a), and the searching sequence of (a) is shown in Figure 3.1(b).



(a)

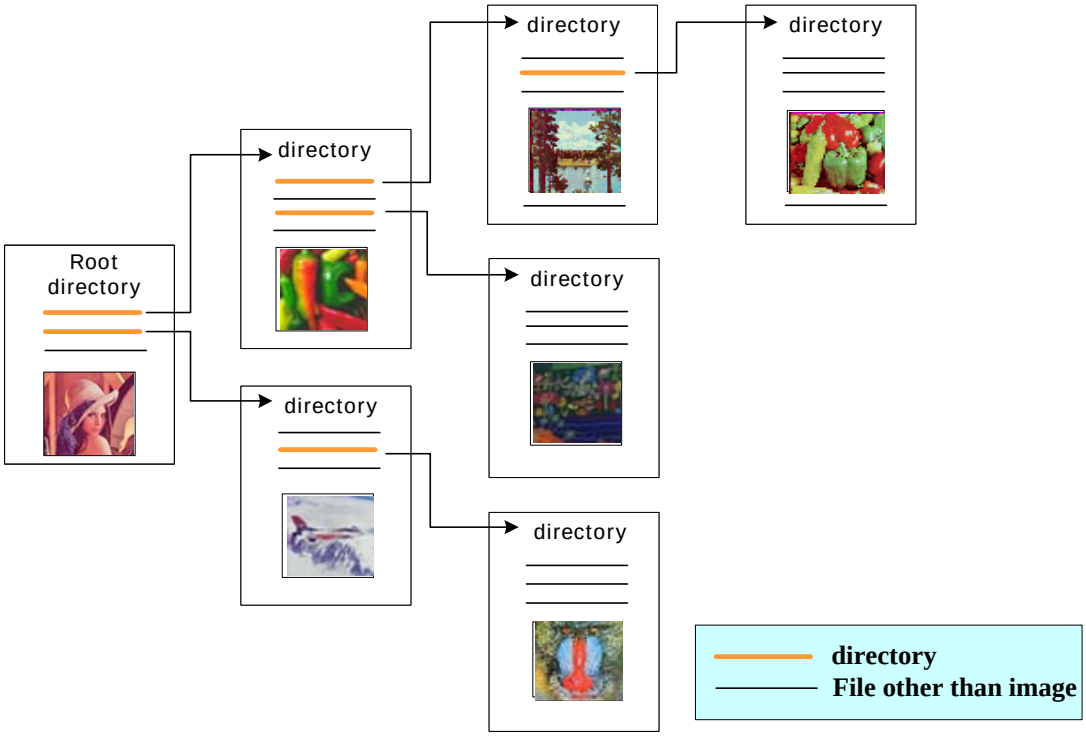
(b)

Figure 3.1 (a) An example of detecting images in webpages within fourth layer using a depth-first search algorithm. (b) Searching sequence of (a).

3.2.2 Proposed Method for Detecting Images in Public FTP Servers

To find suspected images in a public FTP server, we can search all images in the server by parsing the directories of the server. We go through every directory in the server and download images when we find BMP, GIF, or JPEG files in the directory. If there is another directory in this directory, we parse that directory further until every directory has been dealt with. And the method we propose for detecting all

images in public FTP servers is a breadth-first search algorithm. An example of detecting images in a public FTP server using the breadth-first search algorithm is shown in Figure 3.2(a), and the searching sequence of (a) is shown in Figure 3.2(b).



(a)

(b)

Figure 3.2 (a) An example of detecting images in a public FTP server using a breadth-first search algorithm. (b) Searching sequence of (a).

3.2.3 Process for Verification of Watermarked Images

After downloading all suspected images, subsequently we verify the ownership of those images using an applicable watermarking technique. Two watermarking techniques for BMP files are adopted, namely, an invisible watermarking technique and a visible watermarking one. We do the same for GIF files and JPEG files.

At first, we check if there is an embedded signal in the suspected image by an

adopted extraction method. If the suspected image has not been processed by the software InfoProtector before, then we cannot extract any signal from the suspected image. And we judge this suspected image as a non-infringed one. But if the suspected image has been processed by the software InfoProtector before, it will be able to extract the embedded signal from the suspected image. Then we extract all information embedded in the image further to check if the embedded owner is the local image authentication center or the embedded watermark is the same as that registered at the local image authentication center. If the answer is yes, the copyright infringement of the suspected image is confirmed. Otherwise, the suspected image is a non-infringed one. A flowchart of proposed method for verification of watermarked images is shown in Figure 3.3.

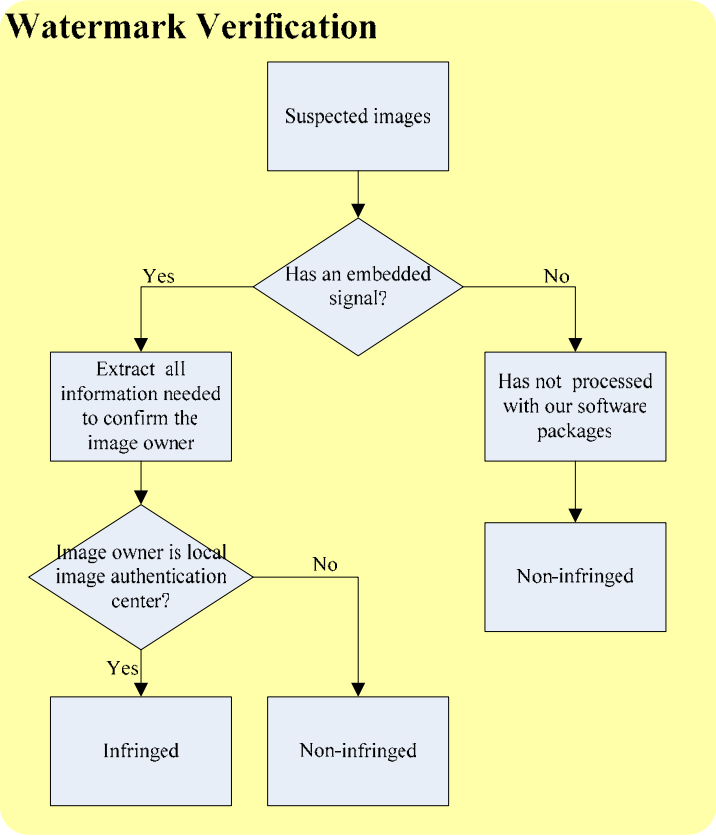


Figure 3.3 A flowchart of proposed method for verification of watermarked images.

3.2.4 Overall Procedure for Proposed Online Search and Verification Method

The procedure for online search and verification of watermarked images for copyright infringement detection has been described in the last two sections, respectively. They are fundamental functions in the proposed system and should always be performed in succession. We can search images in webpages and in public FTP servers by different methods, and then verify the copyright of the acquired images.

Local image authentication centers can do these parts by themselves before requesting the central image authentication center for issuing certificates, or just give the URL of the suspected images and request the central image authentication center for doing these tasks and issuing authentication results. It depends on the loading of the network and the computing power of the central image authentication center. And a person can use a simplified software package, which includes functions of search and verification of copyrighted-infringed images, to prevent misuse of copyrighted images.

A flowchart of the proposed method is shown in Figure 3.4, which represents the overall procedure for online search and verification of watermarked images for copyright infringement detection described previously.

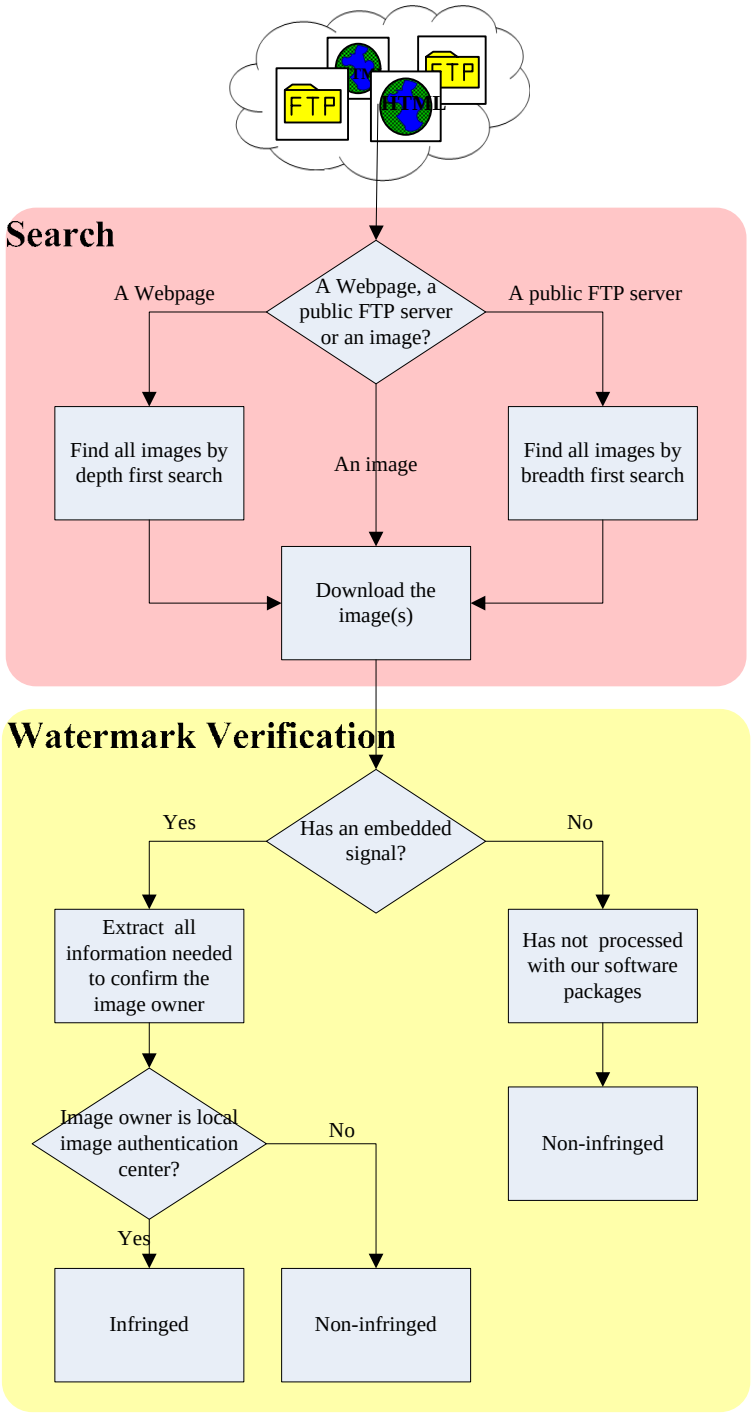


Figure 3.4 A flowchart of overall procedure for online search and verification of watermarked images for copyright infringement detection.

3.3 Discussions and Summary

Method for online search and verification of watermarked images is proposed in this chapter. When we find a suspected image in a webpages or in a public FTP server, we can just download the image or detect all images in the webpages or in the public FTP server by a depth-first search algorithm or a breadth-first search algorithm, respectively. After that, we check whether there is an embedded signal in each image, and then extract all information embedded in images which has been processed by the software InfoProtector before. And we check further if the embedded owner is the local image authentication center or if the embedded watermark has been registered at the local image authentication center. If the answer is yes, the copyright infringement of the suspected image is confirmed. Otherwise, the suspected image is decided to be a non-infringed one.

In addition, the search and verification functions in the proposed system are designed in a passive way. That is, we do not search and verify images until we find a suspected one. Alternatively, an active mode can be implemented in the system. That is, whenever the web browser is on the Internet, it will be searching and verifying watermarked images continuously. But this will result possibly in heavy loading of the network and large consumption of the computing power of the computer. It is very likely that the computer cannot do anything else when the web browser is on the Internet. Thus, we choose to implement our search and verification functions in a passive way.