

要 旨

日本語文書への秘匿情報埋め込み方式

河村 智

近年、コンピュータネットワークの発達とあらゆるコンテンツのデジタル化にともない、誰もが欲しい情報を迅速かつ容易に手に入れることが可能になった。しかし、それに伴い、不正なコピーによる著作権侵害の被害が深刻化してきている。この不正コピーを抑止するためには、模倣された作品が自分のものであることを示し、著作権の侵害を立証しなければならない。

そこで、本稿ではコンテンツ、特に日本語のテキストに透かしを埋め込む手法について提案する。本方式では助詞の同義語置換を行うことによってテキストにビットデータを埋め込む。まず、あらかじめ置換可能な表現をいくつも抽出しておき、さらにどちらがビット0でどちらがビット1に対応するのかを定め、それらを埋め込み規則としてデータベース化しておく。次に、埋め込み手順としては、まずキャリアとなるテキストを文法的に解析し、助詞部分を抜き出す。そして、抜き出した表現を先ほどの辞書データベース内のものと比較し、交換可能な個所があれば候補情報を付加する。次に埋め込みたい透かし情報を2進符号化し、先ほどの候補情報の中からビットデータにあったものを選択し置換を行う。これら一連の処理を繰り返すことによって透かし情報のビットデータが埋め込まれたテキストが完成する。さらに、これらの手順に従い日本語のテキストを解析して、予め用意した辞書データベースに基づいて候補情報を付加し置換を行うプログラムを試作し、本方式の実装を行った。本方式を新聞記事、雑誌のコラム、小説、日記に適用した結果、テキストの内容に左右されることなく、安定した埋め込みを行えることが確認できた。

キーワード 電子透かし、埋め込み、インフォメーションハイディング

Abstract

A Digital Watermarking System onto Japanese Documents

Satoshi KAWAMURA

In recent years, it enabled everyone to get information needed quickly and easily with a spread of a computer network, and digitization of all contents. Therefore, the damage of the literary piracy by the illegal copy is becoming serious simultaneously with it. It proves that the plagiarized work is its thing in order to deter an illegal copy. Furthermore, we must prove inforingement of copyright with it.

Then, I propose a system that embeds a digital watermark in Japanese text. In this system, bit data embedded by replacing the particle of the same meaning. First, much expression of the particle whose replacement is possible is prepared. And it is decided which corresponds to bit 0 in bit 1. And they are registered with a dictionary.

Next, a part of speech of the career text is analyzed and a particle is picked up. Then, the particle refers to the dictionary. And information of the candidate is given to expression of the particle whose replacement is possible. Next, information of the watermark is formed into bit data and the candidates is decided that is becomes same as bit data. Finally, the text in which bit data was embedded by these way. I created the program which embeds bit data in a Japanese text by these way and mounted this system.

key words Watermarking, Embedding, Information Hiding