

# Copyright Protection Method of Big Data Based on Color Image Watermark

Yabin Xu<sup>1</sup>, Bin Shi<sup>2</sup>

1. Beijing Key Laboratory of Internet Culture and Digital Dissemination Research, Beijing 100101, China
2. Beijing Advanced Innovation Center for Materials Genome Engineering, Beijing Information Science and Technology University
3. School of Computer, Beijing Information Science & Technology University, Beijing 100101, China

**Abstract:** Data watermarking technology is an effective means of copyright protection of big data. In order to embed more real and effective color image watermarking information, firstly, based on JPEG image coding standard algorithm, the color image is compressed without affecting the image quality, so as to reduce the embedded watermark information. Then, the watermark is embedded in two dimensions: data tuple and attribute. Finally, under the constraint of data usability, the lowest bit of watermark embedding is calculated and the data is changed to complete the watermark bit embedding. In order to verify the copyright ownership of big data, this paper also presents the corresponding watermark extraction method. Watermark extraction is the reverse process of watermark embedding. First, traverse all tuples and attributes to extract the possible embedded bit values in each attribute element. Then, the actual embedded watermark bit string is determined by majority voting strategy, and the color image is restored after decoding the watermark bit string. Experimental results show that our proposed method can extract color image watermarking information under different attack conditions, the robustness of watermark is high, and the comprehensive effect of data watermark is better than the existing methods.

**Key words:** big data; data watermarking; copyright protection; JPEG image encoding; majority voting strate

