Abstract

The Internet is a widely used paradigm where sharing of multimedia content is a major task. Spam image (an image that contains obscure or irrelevant content) is often discovered in web data available in servers and worldwide search engines. Techniques for spam filtering and finding or detecting obscure content in multimedia data (such as .JPEG, .png format of images) are available in the literature. This paper reviews different existing techniques to deal with obscure images and presents an enhanced ReP-ETD (Repetitive Pre-processing technique for Embedded Text Detection) technique in order to detect the obscured content in image data. The technique proposed in this paper first pre-process the multimedia image data using a Linux image script and further on OCR (Optical Character Reader) is used for the spamming image detection and depth analysis. The main contribution of this paper is to discover and perform spam word extraction from the embedded obscured image.

References
11. Asha S Manek, Shamini D.k, Bhat and Shenoye” ReP-ETD: A Repetitive Preprocessing Technique for Embedded Text Detection from Images in Spam Emails’978-1-4799-2572-8/14/$31.00c@2014 IEEE
Index Terms

Computer Science

Information Sciences

Keywords

K OCR, obscure images, CAPTCHA