IEEE ISIE 2006, July 9-12, 2006, Montréal, Québec, Canada

A Data Hiding Technique for JPEG Color Images by One-Dimensional Spectrum Modification

Kaliappan Gopalan
Department of Electrical and Computer Engineering
Purdue University Calumet
Hammond, IN 46323
E-mail: gopalan@calumet.purdue.edu

***Abstract*-A method of hiding data in images for applications such as carrying identification and authorization information in a picture or logo is proposed. By converting a JPEG-compressed color image to a one-dimensional signal in R, G, or B and using an arbitrary sampling rate of 16 kHz, masked frequencies in the 1-D signal are determined for each segment or block. Imperceptible hiding of given data is carried out by modifying the spectral power at a pair of commonly occurring masked frequencies. Preliminary results show that the spectrum modification technique is simple to process with barely noticeable distortion in the data-embedded image. Using an oblivious technique and a key consisting of the frequencies where spectrum is modified, successful data retrieval with no bit errors has been observed. Embedded image corrupted by low level noise still retained the hidden data with low bit errors.**

I. INTRODUCTION

Data embedding employing a host image is a useful means for storing information with increased storage capacity and for covert transmission or steganography. Information is hidden in a cover image in such a way that the embedded image (the stego) is indiscernible from the unembedded host, or cover image. By concealing information imperceptibly and using a strong key, attempts at illegal recovery and use of sensitive data are foiled. An imperceptible embedding technique that can also recover the hidden information accurately and without requiring the cover image, i.e., by an oblivious method, can be used in secure communication [1]. Use of an oblivious data recovery technique enables deploying any innocuous cover image for carrying battlefield information in covert or secure communication.

Another key application area is in embedding vital medical and biometric information of employees in their pictures for secure identification and/or data retrieval [2, 3]. In these applications, biometrics-based identifying information, for example, may be hidden in the picture card of a person and the claimed identity of the card carrier can be verified by retrieving the hidden data; as such, a small amount of distortion in image quality may be tolerable as long as data robustness is assured. Watermarking of images and video for compliance of digital rights management applications, on the other hand, requires spreading of a small amount of data. While imperceptibility and payload are critical for covert communication, data robustness is vital in personnel authentication and watermarking applications with, preferably, oblivious data extraction.

We present a method of embedding data in a color image that requires a strong key for retrieval using spectral domain modification. Image distortion, hence perceptibility of embedding, is minimized at a cost of lower payload, with the embedded information spread throughout the image. This method is proposed as an extension to prior work on spectral domain audio embedding by tone addition [4-6] and gray level image embedding [7].

II. SPECTRAL DOMAIN EMBEDDING

Imperceptibility of embedding data in an image can be achieved by exploiting the imperfection in the human visual system, which has lower detectability of patterns in the presence of other patterns at close frequencies. Based on the results of secure embedding in the spectral domain of audio signals, the proposed technique for image embedding relies indirectly on the masking property of the human visual system. In the case of audio embedding at psychoacoustically masked audio frequencies, a two step procedure is used [4]. In the first, a set of auditorily masked spectral points for each segment (frame) of the cover audio signal is determined. These frequencies for each short segment of speech depend on the just noticeable difference (JND) in hearing and a global masking threshold based on a set of critical band filters. A pair of masked frequencies that occur in most segments of a given cover speech signal are obtained using a common minimum threshold of sound pressure level (SPL) relative to the global masking threshold of each segment. Since the two frequencies in each segment have SPL below the hearing (global) threshold, an increase in SPL up to the global threshold level (or decrease from its current level) cannot alter hearing perception. Based on this premise, SPL at the two frequencies can be set to a known ratio in accordance with binary data to be embedded.

Modification of the spectrum is carried out in the second step by setting the power levels at the two masked frequencies in a known ratio in accordance with the bit value to be embedded in the frame. The pair of frequencies and the power ratio of the two masked spectral components form the key for embedding and retrieval of data. In the case of audio embedding, average power levels set to one-tenth and one-hundredth of the segment power at masked frequencies have been observed to result in inaudible and robust hiding of information. Since the modification to the two spectral points are at relatively low power levels, and it is carried out in the spectral domain, the resulting spectral change is spread across all audio samples in a segment, thereby rendering the embedded (stego) audio indistinguishable from the original (cover) audio. Additionally, spreading of embedding makes detection by steganalysis difficult to accomplish.

Extension of spectrum modification at selected frequencies for image embedding and its implementation are described in the following sections.

III. IMAGE SPECTRUM MODIFICATION PROCEDURE

To extend the above two-step audio embedding algorithm for hiding data in an image, a common pair of visually masked spectral points can be determined using psychovisual contrast or pattern masking frequencies from the discrete cosine transform (DCT) of each block of an image [8, 9]. Alternatively, a simpler one-dimensional approach, similar to the determination of psychoacoustically masked frequencies for an audio signal, can be used in place of detecting the JND in the image. This approach can be further simplified by converting the two-dimensional intensity level of a color host image in one of the three primary colors (or, the gray level of a black-and-white host image) to one-dimensional signal by appending all the rows (or columns) sequentially. (It has been shown that by treating each block of 8x8 subimage as a frame (by conversion to one-dimensional data) of ‘audio’ and appending the blocks together causes noticeable distortion in the embedded image [7]).

Choosing a high enough ‘sampling frequency,’ a pair of most commonly occurring masked frequencies in all frames (of typically 64 pixel samples each) are obtained by determining global masking threshold and setting an acceptable level of spectral density below this level for each frame [10]. (Although the choice of sampling frequency is empirical, a high value – above 10,000 Hz – gives more masked frequencies which contribute to a stronger key for embedding and retrieval.) Spectral power levels at the selected pair of frequencies in each frame (or, for added security, in selected frames) are set by a known ratio for embedding binary values. An advantage of this conversion and embedding is that it entails less computational complexity and faster detection of embedding points compared to a DCT-based procedure. Since there is no relationship between the audibility of a masked tone and the visibility of a masked pixel, however, there is no guarantee that an audibly masked frequency will result in a visibly masked two-dimensional frequency pair. Still, the availability of a set of audibly masked frequencies offers a choice for spectrum modification.

A pair of most commonly occurring masked frequencies *f1* and *f2* form the key for embedding and retrieval. Complex spectrum at each of the two frequencies is modified to attain imperceptibility of embedding. Since the two audibly masked frequencies may not be present in all the segments, raising their power levels based on the global audio masking threshold for a given frame may result in discernibility of embedding in the overall audio and, hence, image. To prevent this, power levels at *f1* and *f2* are set to low levels in each segment as follows.

To embed a 1:

$$X'(f1)=\alpha.e^{i\theta_1}$$
$$X'(f2)=\beta.e^{i\theta_2} \tag{1a}$$

To embed a 0:

$$X'(f2)=\alpha.e^{i\theta_2}$$
$$X'(f1)=\beta.e^{i\theta_1} \tag{1b}$$

where $X'(f1)$ and $X'(f2)$ are the modified spectral components at frequencies *f1* and *f2*, and θ_1 and θ_2 are the phase angles of the original spectrum at *f1* and *f2*. The constants α and β are adapted based on the average power of each segment. Typically, α is larger than β so that the spectrum at one of the two frequencies is higher than that at the other frequency. Both values, however, are small enough so that they are not visible in the spectrogram of the audio signal and large enough to be not lost in quantization after embedding.

Since the two frequencies *f1* and *f2* for an embeddable frame, i.e., one selected for spectrum modification, are in its masked region, adding or subtracting spectra at these frequencies ensures that the modification results in minimal change in ‘audibility.’ Also, by retaining the same phase as that in the original spectrum at *f1* and *f2*, no phase-related distortion is introduced after modification. Modified frame (block) spectrum is transformed to time domain, quantized to the same number of levels as the cover image, and converted back to two-dimensional image.

Embedded information in each segment is recovered by the spectral ratio at the two (key) frequencies, *f1* and *f2*. That is, the recovered bit *rb* in a segment (image block) is given by

$$rb = \left\{ \begin{array}{l} 1, \text{ if } \left| \dfrac{X(f1)}{X(f2)} \right| > b1 \\ 0, \text{ if } \left| \dfrac{X(f2)}{X(f1)} \right| > b0 \end{array} \right\} , \tag{2}$$

where *X(f)* is the spectral component of the frame (block) at cyclic frequency *f*, and *b1* and *b2* are set empirically. If a segment is left unembedded for added security or when the data size is smaller than the available capacity, spectral levels at the two frequencies are set equal; this corresponds to a retrieved bit of -1.

The key for embedding and retrieval consist of the indices of the embedded frames (image blocks), if only a selected frames are used to carry hidden data, and the corresponding frequency pair used to modify spectrum. This key, clearly, depends on the cover (host) image. Additionally, a given cover image may have more than one pair of embeddable (masked) frequencies. Both the variability and the presence of many masked points make it harder for illegal retrieval and/or tampering of data by an exhaustive search of possible embedded frequencies.

IV. IMPLEMENTATION AND RESULTS

Results of the two-step image embedding algorithm using the gray scale cameraman image available in Matlab showed that embedding at a pair of high frequencies, even though they are not the most commonly occurring masked frequencies, caused little noticeable distortion and zero bit error in data recovery. Based on these results, the technique was applied to embedding data in one or more of the primary colors in a JPEG color image. This image (kid.jpg) of 289x200x3 pixels was converted one-dimensional signal in each of the three colors by appending all the rows of pixel values together. Using an arbitrary sampling frequency of 16,000 Hz, the masked frequencies for each color were obtained. The pair of frequencies, 4875 Hz and 6250 Hz, which were in the masking set of fewer than 30 percent of the segments for all three colors, resulted in minimal distortion of embedded image with each of the three colors. For the size of 289x200 = 57800 values in the one-dimensional signal, a maximum of 57800/64 = 903 bits can be embedded when all the frames of 64 pixels are used.

To test the image quality with this full capacity, (a) all bits of 1, (b) all bits of 0, and (c) a random order of binary values were used for the data with the constants α and β set at a ratio of 1E-5 from the average power of each segment. The resulting image for (a) is shown in Fig. 1 along with the original cover image. Using a spectral ratio of $b1 = b2 = 1$ in Eq. (2), all the embedded bits were retrieved correctly from the modified and quantized image. Perceptibility of embedding, as can be seen from the figure, appears to be minimal. A pixel-by-pixel difference in the blue color array that had been changed, showed a range of [-66 84] with the blue array of the original host image. Similar results in data recovery and perceptibility were observed for the other two cases as well with all 903 bits of 1's or 0's. Fig. 2 shows two cases for hiding in red and green colors.

Extending the embedding procedure to more than one color using the same pair frequencies showed a slightly noticeable distortion in the image, as can be seen in Fig. 3. Considering that the payload is doubled or tripled, the distortion may be tolerable in pictures such as in a driver license, for example. Although use of different frequencies for each color altered perceptibility of embedding slightly, the key became stronger with three pairs of frequencies.

Original

(a)

Stego, Blue with all 1's

(b)

Fig. 1. (a) Original host image, and (b) Image with blue modified at 903 pixels

Preliminary results on the robustness of data with additive noise showed low bit-error rate at low noise power levels. Further work on robustness is in progress.

Stego, Red with all 1's

(a)

Stego, Green with all 0's

(b)

Fig. 2. Hiding (a) 903 bits of 1's in red, and (b) 903 bits of 0's in green

Stego, R & G

Fig. 3. Hiding in red and green with a total of 1806 bits

V. Conclusion

A method of embedding data on a JPEG color image by converting the image to a one-dimensional signal in each color has been proposed. By altering the one-dimensional spectrum of each segment of a cover image at two key frequencies, embedding becomes barely noticeable. Availability of a choice of frequencies for the key at which the spectrum is modified renders the hidden data impervious to unauthorized access. Another advantage of the technique is that the hidden information is extracted by an oblivious method. The method can be used to embed authorization information in the picture of an employee in access control applications.

Embedded data can be made resistant to illegal access by using a different frequency pair for each block. Payload or hidden data capacity can possibly be increased by using four frequencies to modify the spectrum, and/or using more than one color at a cost of image distortion.

A key question that arises from the proposed method is the lack of correlation between audibly masked frequencies and the JND in each image frame. Another is the choice of an appropriate sampling frequency in the conversion so that an embedded image is indistinguishable from its original cover image. Since there is no relationship between the audibility of a masked tone frequency and the visibility of a masked pixel, the implicit assumption in going from one-dimensional (audible) to two-dimensional (visible) domain may not always result in imperceptible embedding. The simplicity of the proposed technique, therefore, must be weighed against these questions. By empirically evaluating the masked frequencies for a choice of one-dimensional sampling frequencies, it is expected that the proposed technique can be applied with minimal visibility of embedding.

References

[1] F.A. P. Petitcolas, R.J. Anderson and M.G. Kuhn, Information hiding – a survey," *Proc. IEEE*, Vol. 87, No. 7, pp. 1062-1078, 1999.

[2] M. Wu and B. Liu, “Data hiding in image and video .I. Fundamental issues and solutions,” *IEEE Transactions on Image Processing*, Vol. 12 , pp. 685 – 695, June 2003.
[3] M. Wu, H. Yu and B. Liu, “Data hiding in image and video .II. Fundamental issues and solutions,” *IEEE Transactions on Image Processing*, Vol. 12 , pp. 696 – 705, June 2003.
[4] K. Gopalan, S. Wenndt, S. Adams and D. Haddad, “Covert Speech Communication Via Cover Speech By Tone Insertion,” *Proc. of the IEEE Aerospace Conference*, Big Sky, MT, Mar. 2003.
[5] K. Gopalan and S. Wenndt, “Audio Steganography for Covert Data Transmission by Imperceptible Tone Insertion,” *Proc. the IASTED International Conference on Communication Systems and Applications (CSA 2004)*, Banff, Canada, July, 2004.
[6] K. Gopalan, “Audio Steganography by Cepstrum Modification,” *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005)*, Philadelphia, March 2005.
[7] K. Gopalan, “A Spectrum Modification Technique for Embedding Data in Images,” *Proc.* the *38th Southeastern Symposium on System Theory*, Cookeville, TN, March 2006.
[8] B. Wolfgang, C.I. Podilchuk, and E.J. Delp, “Perceptual watermarks for digital images and video,” *Proc. IEEE*, Volume: 87, Issue: 7, pp.1108 – 1126 July 1999.
[9] W. Bender, D. Gruhl, N. Morimoto and A.Lu, “Techniques for data hiding,” *IBM Systems Journal*, Vol. 35, Nos. 3 & 4, pp. 313-336, 1996.
[10] T. Painter and A. Spanias, “Perceptual coding of digital audio,” *Proc. IEEE*, vol. 88, No. 4, pp. 451-513, Apr. 2000.