Attack By Colorization of a Grey-Level Image Hiding its Color Palette

Marc Chaumont, William Puech

ATTACK BY COLORIZATION OF A GREY-LEVEL IMAGE HIDING ITS COLOR PALETTE

Chaumont M. and Puech W.

Laboratory LIRMM, UMR CNRS 5506, University of Montpellier II
161, rue Ada, 34392 MONTPELLIER CEDEX 05, FRANCE

ABSTRACT

In this paper, we present a novel attack named *colorization attack*. This attack is specific to color-hiding watermarking schemes. The objective of this work is to demonstrate the feasibility of such an attack and thus to take it into account for the future color-hiding watermarking schemes.

Index Terms— watermarking, color-hiding, colorization attack, cost-function optimization

1. INTRODUCTION

The colorization is the act of adding color to a monochrome image or a movie [1]. This term was introduced by Wilson Markle in 1970 to describe its method for adding color to black and white movies or TV programs. Various techniques have been proposed [2, 3, 4], most of them based on an image segmentation into regions. Those techniques necessitate lots of manual intervention in order: to attribute a color to each regions and to correct the wrongly segmented regions. Recently, the approach of Levin *et al.* [5] based on a very simple energetic model and its optimization leads to amazing good colorization results. Moreover this approach necessitates a very small manual intervention. In this paper, we propose to study the *colorization attack*. This attack is specific to watermarking schemes **hiding the color information** and is an indirect attack in order to retrieve a color image closed to the original one.

Few solutions have been recently proposed in order to **protect** the **color information** with the **data-hiding** paradigm. The grey-level image (embedding the color information) is freely accessible but its color version necessitates to own a key. The first group of solutions is based on wavelet decomposition and sub-band substitution and embed the chroma information in a grey-level image [6, 7, 8]. The purpose is not specifically to protect the color information but more classically to propose perceptive compression and image authentication for Campisi *et al.* [6] and Zhao *et al.* [7] and printing solution for Queiroz and Braun [8]. The second group of solutions is based on the decomposition of a color image into an *index* image and its color palette then the embedding of the color palette image into the *index* image [9, 10].

All those color-hiding solutions produce a grey-level image embedding the color information. The **hidden** color information is used to rebuild the color image. For all those schemes, the **enriched** grey-level **image** is very close to the original luminance image. A **colorization attack** may then be an **indirect** attack to retrieve **a** color image visually pleasant. Even if the colorized image is far from the original color one, it could give an attractive color version and then yield null and void the color-hiding watermarking schemes. Indeed, the added value of the color protection schemes is the color information; if a colorized and nice version is easily generated, the watermarking schemes are no more reliable because, for example, an illegal color printing may be sold.

In this paper we specially treat of colorization attack for the palette-based watermarking scheme [9, 10]. The subband wavelet substitution watermarking scheme [6, 7, 8] may them been attacked thanks to the colorization exposed in [5] or with an adaptation of our proposed method. Comparing to [5], we express analytically the solution to the problem (authors of [5] use a full "matrix inversion" and do not describe explicitly initials conditions) which involves, for the colorization process, a CPU complexity reduction . Moreover, we adapt the *colorization attack* to the palette-based watermarking schemes which implies an easier initialization: user gives generally less than 7 points with their associated colors.

2. THE COLORIZATION ATTACK

2.1. General colorization formulation

As shown in [5] a grey-level image may most of the time be nicely colorized with small human intervention and in low CPU complexity. In this section, we propose a scheme in order to colorize grey-level images that are in the same time $index$ image ("colorization attack" of palette-based schemes). The palette-based schemes own a strong property: there is a bijection between grey-level values and a colors. It is thus easier to found just K colors compared to [5] where the number of unknowns equals to the size of the image.

As we already own an intensity information (the *index* image) we just need to extract the two chrominances planes. Thus, we decide to work in the YUV color space where the luminance plane Y is known (it is the *index* image) and where

U and V are the **two unknown chrominance planes**. Moreover, finding this two unknown planes is equivalent (in the palette-based watermarking scheme) to find the K colors of the palette C. Our unknown is thus the color palette C.

The model for the colorization problem is very simple: two spatially close pixels i and j should own a similar color if their intensity are similar. Mathematically, this may be expressed as minimizing for a pixel i all **weighted neighborhood differences** between the color $C(Index(i))$ of pixel i and the color $C(Index(j))$ of the pixel j belonging to the neighborhood of i which is noted $\mathcal{N}(i)$:

$$\forall i, \sum_{j \in \mathcal{N}(i)} w_{i,j}.(C(Index(i)) - C(Index(j)))^2, \quad (1)$$

where $w_{i,j}$ is a weighting function, large when $Index(i)$ is similar to $Index(j)$ and small when this two intensities are different. The commonly used weighting function by segmentation algorithms for measuring the similarity between two intensity $Index(i)$ and $Index(j)$ is the Gaussian function:

$$w_{i,j} = a.e^{\frac{-(Index(i)-Index(j))^2}{2\sigma_i^2}},$$

with a, a positive real constant and σ_i^2 the local variance in a window around pixel i. Note that those weighting functions are unsymmetric ($w_{i,j}$ is not necessary equal to $w_{j,i}$). Also note that other weighting functions may be used as proposed in [5]. Those weighting functions give a value around 1 if $Index(i)$ and $Index(j)$ are close with respect to the local variance σ_i^2 and a value near 0 otherwise.

2.2. Specific colorization formulation

Note that the cost function E objective in [5] is to minimize the difference between the color of pixel i and the **weighted neighborhood colors**:

$$E(x) = \sum_{i=1}^{i=N} \left(x(i) - \sum_{j \in \mathcal{N}(i)} w_{i,j} x(j) \right)^2, \quad (2)$$

where x is either the U plane (of size N pixels), either the V plane (of size N pixels) of the color space YUV. The objective is to find the two unknown U and V. This energetic model owns the same form than the one used in segmentation algorithms based on normalized graph cuts [11]. The solution is obtained by computing the second smallest eigenvector of $D-W$ with W the $N \times N$ weighting matrix and D the diagonal matrix [5]. The minimization is simply proceeded trough eigenvector decomposition of the $N \times N$ high dimension matrix $D - W$.

With the specific case of palette-based watermarking schemes the cost function may be re-written:

$$E(C) = \sum_{i=1}^{i=N} \left(C(Index(i)) - \sum_{j \in \mathcal{N}(i)} w_{i,j}.C(Index(j)) \right)^2, \quad (3)$$

where C is the unknown color palette. The equation form is not similar to the previous one (2) and we could not use the eigenvector decomposition approach. We then decide to re-write a close form for equation (3) in order to facilitate its analysis and to found an analytic solution for the optimization. We are moreover adding an initialization constraint for user colors guess. The user manually sets L couples (i_l, c_l) where i_l is a pixel position and c_l its associated color. In order to respect the user choices we impose that the color $C(Index(i_l))$ of the pixel i_l will be a color close to given color c_l i.e. minimize $(C(Index(i_l)) - c_l)^2$. The cost function is thus made of a constraint term on neighborhood differences (equation 1) and a constraint term on user colors guess:

$$\begin{aligned} E(C) &= \sum_{i=1}^{N} \sum_{j \in \mathcal{N}(i)} w_{i,j}.(C(Index(i)) - C(Index(j)))^2 \\ &\quad + \lambda \sum_{l=1}^{L} (C(Index(i_l)) - c_l)^2. \end{aligned} \quad (4)$$

The solution of equation (4) is obtained by canceling $\frac{\partial E}{\partial C(k)}$. The algorithm is iterative and for each k, the $C(k)$ colors (two chrominances U and V unknown) are updated until convergence such that (detail algorithm is given on Listing 1): $\forall k \in [1, K]$,

$$\begin{aligned} N(k) &= \sum_{i|Index(i)=k} \left(\sum_{j \in \mathcal{N}(i) \text{ and } Index(j) \neq k} w_{i,j}.C(Index(j)) \right) \\ &\quad + \sum_{i|Index(i)\neq k} \left(\sum_{j \in \mathcal{N}(i) \text{ and } Index(j) = k} w_{i,j}.C(Index(i)) \right) \\ D(k) &= \sum_{i|Index(i)=k} \left(\sum_{j \in \mathcal{N}(i) \text{ and } Index(j) \neq k} w_{i,j} \right) \\ &\quad + \sum_{i|Index(i)\neq k} \left(\sum_{j \in \mathcal{N}(i) \text{ and } Index(j) = k} w_{i,j} \right) \\ C(k) &= \begin{cases} \frac{\lambda.c_l + N(k)}{\lambda + D(k)} & if\ \exists l,\ Index(i_l) = k, \\ \frac{N(k)}{D(k)} & if\ \nexists l,\ Index(i_l) = k \end{cases} \end{aligned} \quad (5)$$

3. RESULTS AND DISCUSSION

Two *colorization attacks* have been proceeded on the color-hiding scheme of Chaumont and Puech [10]. The pirate was familiar to signal processing and had only access to the watermark image (i.e the *index* image). The two concern images are *baboon* 256×256 and *kodak-13 mountain stream* 256×384.

Figures 1.b and 2.b show colorization results obtained respectively with 5 and 6 user colors. Note that the user has been asked to choose few colors without any knowledge of the original color version. Figures 1.c and 2.c give the rebuilt

color images knowing the decoding key i.e knowing the hidden color palettes. One can note a difference between the colorized images and the rebuild images or between the palettes obtain by a colorization attack (Figures 1.e and 2.e) and the hidden palettes (Figures 1.f and 2.f).

Nevertheless, remember that the attack should produce a pleasant color version which could be print and sold. In those conditions, the visual quality of the obtained colorized images are very good. Also note that it was not asked to the user to spend time to tune its results. Those results are obtained in less than 5 tries by image which take less than 5 minutes. Improved results may have been obtained with more allocated time and with an adapted visual interface.

As it is clear that colorization attack may be potentially a security hole; counter-attack should be proposed in order to improved color-hiding watermarking schemes. An immediate solution for those watermarking scheme is to generate a watermark grey-level image which is far from the original luminance image. Other propositions have to be proposed in the future in order to improve color-hiding schemes.

4. CONCLUSION

To conclude, this paper introduces a new (indirect) attack specific to *color-hiding schemes*. This attack does not allow to retrieve the original color version of a watermark image, nevertheless a pleasant colorized image may easily be created. Future *color-hiding schemes* should then take into account this particular attack in order to be more robust.

5. REFERENCES

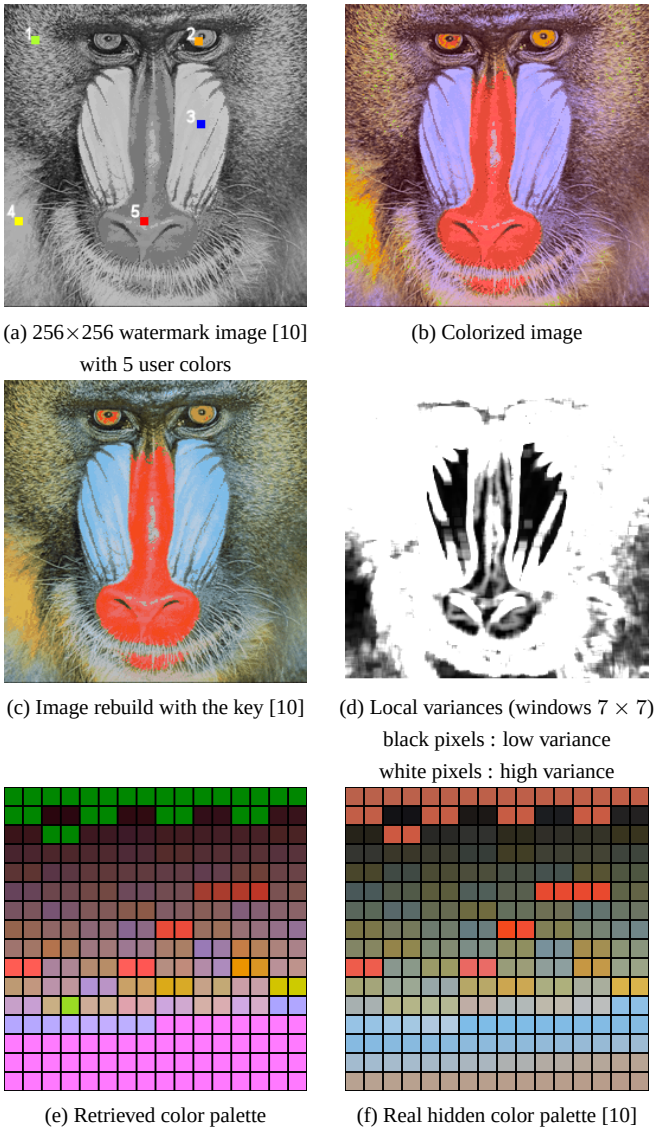
[1] G. Burns, *Colorization (The Encyclopedia of Television)*, Museum of Broadcast Communications, http://www.museum.tv/archives/etv/index.html.

[2] W. Markle and B. Hunt, "Coloring a Black and White Signal Using Motion Detection," in *Canadian patent no. 1291260*, Dec. 1987.

[3] J. Silberg, "The Pleasantville Post Production Team that Focussed on the Absence of Color," Cinesite Press Article, http://www.cinesite.com/core/press/articles/1998/10 00 98-team.html, Oct. 1998.

[4] NeuralTek, "BlackMagic Photo Colorization Software," 2003, http://www.timebrush.com/blackmagic.

[5] A. Levin, D. Lischinski, and Y. Weiss, "Colorization using Optimization," in *International Conference on Computer Graphics and Interactive Techniques, ACM SIGGRAPH'2004*, Los Angeles, California, 2004, pp. 689–694.



(a) 256×256 watermark image [10] with 5 user colors

(b) Colorized image

(c) Image rebuild with the key [10]

(d) Local variances (windows 7×7) black pixels : low variance white pixels : high variance

(e) Retrieved color palette

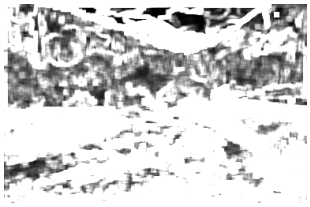
(f) Real hidden color palette [10]

Fig. 1. Illustration of the colorization algorithm on baboon image

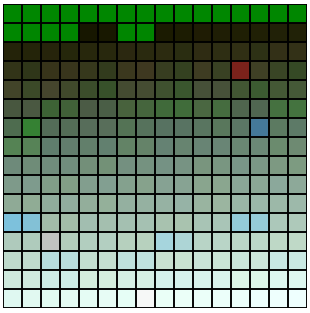
(a) 256×384 watermark image [10] with 6 user colors

(b) Colorized image

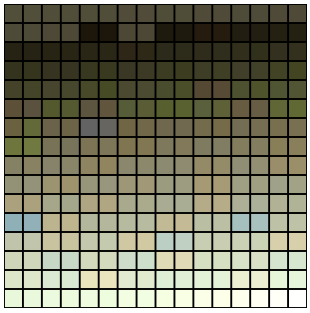
(c) Image rebuild with the key [10]



(d) Local variances (windows 7 × 7) black pixels : low variance white pixels : high variance



(e) Retrieved color palette



(f) Real hidden color palette [10]

Fig. 2. Illustration of the colorization algorithm on kodak-13 *mountain stream* image

Listing 1. *Colorization algorithm*

```
const Integer NBITER; // maximum number of iteration

Procedure Colorization():Palette
begin

  Palette Cnew, Cold; // current and previous color palette

  // TAKE INTO ACCOUNT THE L USER COLORS ;
  // ∀l ∈ [1, L], Cnew(Index(i_l)) ← c_l
  init(Cnew);

  // ITERATIONS
  loop until NBITER reached or CONVERGENCE reached
  begin

    // COPY Cnew INTO Cold
    Cold ← Cnew;

    // UPDATE THE PALETTE Cnew WITH Cold KNOWLEDGE
    Apply equation 5;

    // SET IN CONFORMANCE EACH COLOR Cnew(k)
    // (U and V pixels ∈ [0,255])
    conformance(Cnew);

  end

  // RETURN PALETTE
  return Cnew;
end
```


[6] P. Campisi, D. Kundur, D. Hatzinakos, and A. Neri, "Compressive Data Hiding: An Unconventional Approach for Improved Color Image Coding," *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 2, pp. 152–163, 2002.

[7] Y. Zhao, P. Campisi, and D. Kundur, "Dual Domain for Authentication and Compression of Cultural Heritage Images," *IEEE Transaction on Image Processing*, vol. 13, no. 3, pp. 430–448, 2004.

[8] R. de Queiroz and K. Braun, "Color to Gray and Back: Color Embedding Into Textured Gray Images," *IEEE Transaction on Image Processing*, vol. 15, no. 6, pp. 1464–1470, 2006.

[9] M. Chaumont and W. Puech, "A Fast and Efficient Method to Protect Color Images," in *IS&T/SPIE 19th Annual Symposium on Electronic Imaging, Visual Communications and Image Processing, VCIP2007, SPIE2007*, San Jose, California, USA, Jan. 2007, vol. 6508.

[10] M. Chaumont and W. Puech, "A Grey-Level Image Embedding its Color Palette," in *IEEE International Conference on Image Processing, ICIP'2007*, San Antonio, Texas, USA, Sept. 2007.

[11] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.