Reconhecimento de adulteração de imagens JPEG através da inconsistência do Block Artifact Grid

Lucas Marques da Cunha[1]
Thaís Gaudêncio do Rêgo[2]
Leonardo Vidal Batista[3]

Resumo – A criação e a comercialização de softwares de edição de imagens permitiu que pessoas comuns pudessem realizar qualquer tipo de manipulação em imagens digitais. Em um cenário judicial em que autenticidade e integridade dos dados é crucial, se faz necessário técnicas que permitam garantir tais atributos. A análise forense em imagens digitais busca através de métodos computacionais científicos, reconhecer a presença ou ausência desses atributos. O presente trabalho apresenta um método de reconhecimento de adulteração em imagens JPEG. Esse método baseia-se na técnica de análise da inconsistência do BAG (Block Artifact Grid) da imagem que é gerado a partir de técnicas de adulteração, tais como composição e clonagem. O BAG da imagem trata-se da demarcação dos blocos JPEG. Os testes foram realizados em 60 imagens, utilizando a taxa de acurácia como métrica para a avaliação da efetividade do método.

Palavras-Chave: Integridade, Autenticidade, Reconhecimento, Compressão.

Abstract – The manufacturing and marketing of image editing softwares allowed ordinary people could do any kind of manipulation of digital images. In a judicial setting where authenticity and integrity of data is crucial, some techniques are necessary to ensure such attributes. Digital image's forensic analysis seeks through scientific computational methods to recognize the presence or absence of these attributes. This work presents a tamper recognition method in JPEG images. This method is based on the analysis of inconsistency of image's BAG (Block Grid Artifact) that is generated from tampering techniques, such as composition and cloning. The image's BAG is the demarcation of JPEG blocks. The tests were performed on 60 images using the accuracy rate as an evaluating metric for effectiveness of the method.

Keywords: Integrity, Authenticity, Recognition, Compression.

1. Introdução

Com o grande crescimento da comercialização de câmeras digitais, foram criados vários *softwares* de edição de imagens, tais como *Adobe Photoshop, Gimp, Paint.NET, Corel*

[1] Estudante de Mestrado pelo Programa de Pós-graduação em Informática –PPGI da Universidade Federal da Paraíba – UFPB.
[2] Doutora em Ciência da Computação pela Universidade Federal do Pernambuco – UFPE, professora do Centro de Informática da Universidade Federal da Paraíba – UFPB.
[3] Doutor em Engenharia Elétrica pela Universidade Federal da Paraíba – UFPB, professor do Centro de Informática da Universidade Federal da Paraíba – UFPB.

PaintShop Pro, *Apple Aperture*. Essas edições são realizadas buscando melhorar a qualidade final da imagem, porém, há casos em que estes programas são utilizados de maneira maliciosa, buscando, muitas vezes, modificar o contexto da cena, uma vez que as imagens digitais exercem grande influência na interpretação humana. Dessa forma, há uma grande dificuldade em identificar visualmente uma possível adulteração nas imagens devido à grande eficiência das técnicas existentes. Nesse cenário, existe uma subárea da Computação Forense responsável pela promoção da integridade e autenticidade dessas imagens. Tais atributos são cruciais em processos judiciais e podem ser atestados por meio de técnicas computacionais científicas, desenvolvidas de acordo com cada problema específico.

As modificações em imagens geralmente ocorrem para fins estéticos, em que o usuário busca melhorar visualmente a imagem tornando-a adequada a interpretação humana ou computacional, como aumento de brilho e contraste, redução de ruído, enfoque em objetos da cena, entre outras. Em contrapartida, há alterações para fins maliciosos, em que o objetivo é modificar o aspecto semântico da imagem. Nesse último caso, há diversas técnicas que são utilizadas para esse fim (ROCHA; GOLDENSTEIN, 2010). Algumas dessas técnicas são descritas a seguir:

- **Composição (*splicing*)**: através dessa técnica é possível extrair conteúdo de várias imagens para produzir uma imagem falsa. Utiliza-se um objeto de uma imagem A e introduz-se em uma imagem B. Esse tipo de adulteração pode causar inconsistência de artefatos de compressão, iluminação, padrão CFA e padrão de ruído do dispositivo (BURVIN; ESTHER, 2014).
- **Ajuste fino de Bordas (*feather edges*)**: essa técnica complementa a anterior, dado que após realizada a composição de imagens, ajusta-se as bordas para que haja o mínimo de artefatos introduzidos durante a junção de imagens.
- **Casamento de padrões de iluminação (*light matching*)**: quando há combinação de imagens, geralmente pode ocorrer inconsistência de iluminação, assim, essa técnica permite ajustar a iluminação da imagem combinada, fazendo com que seja imperceptível a mudança, podendo ser distinguível apenas com o uso de um método adequado.
- **Realce:** permite que objetos presentes na cena de uma imagem sejam obscurecidos ou enfatizados. Assim, a partir dessa técnica, aumenta-se a dificuldade de visualizar qualquer modificação feita na imagem.

- **Imagens geradas por computador:** são modelos tridimensionais baseados em imagens reais. São utilizadas principalmente em animações. Podem ser utilizadas técnicas de adição de textura e aplicação de cor de modo que a imagem torne-se o mais próximo da realidade.
- **Cópia e Colagem (*Cloning*):** consiste em retirar partes da própria imagem e adicioná-los em outro local da cena. Esse tipo de adulteração pode causar inconsistência tanto no padrão CFA da imagem quanto nos artefatos de compressão JPEG (FRIDRICH *et. al.*, 2003).

A aplicação conjunta dessas técnicas torna o processo de reconhecimento de adulteração mais difícil, sendo necessária a utilização de métodos combinados para verificação da integridade. A Figura 1 apresenta a aplicação das técnicas de manipulação descritas anteriormente.

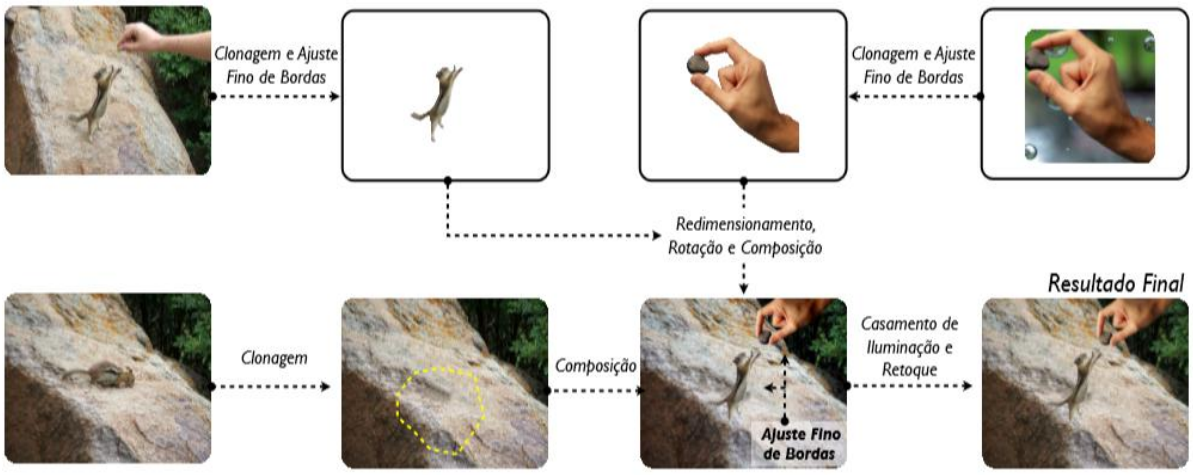

Figura 1. Utilização conjunta das técnicas de manipulação descritas. Nessa exemplo observa-se a combinação de imagens distintas com o intuito de modificar o contexto da cena apresentada por meio das técnicas: Composição, Cópia e Colagem, Ajuste finos das bordas, Ajuste de Iluminação.
Fonte: (ROCHA; GOLDENSTEIN, 2010)

Essas técnicas exploram as diversas particularidades que as imagens possuem, sendo a mais comum, a análise da disposição das cores do sensor de um dispositivo de captura. Assim, de acordo com o modelo do dispositivo, os sensores podem apresentar padrões de cores distintos, variando a partir da sua Matriz de Filtro de Cor (*Color Filter Array*, CFA), sendo que a maioria das câmeras comerciais utiliza o padrão de Bayer. A análise do CFA pode ser utilizada para a autenticação, identificação de adulterações, reconhecimento de

modelos de dispositivos e dispositivos específicos. Nesse contexto, vários autores propõem métodos baseados no CFA da câmera. Popescu e Farid (2005) utilizam EM (Expectation Maximization) para detecção do algoritmo de interpolação CFA e identificação de adulteração em imagens digitais. Bayran *et. al.* (2008) utilizam os algoritmos de interpolação CFA para identificação do modelo do dispositivo de captura, utilizando coeficientes e magnitudes no espectro no domínio da frequência. O método apresentando em Gallagher e Chen (2008) utiliza o padrão CFA para autenticação de imagens e baseia-se na variância de ruídos dos pixels interpolados e não interpolados, assim, verifica a existência ou não de traços de interpolação de forma que seja possível diferenciar imagem real e imagem gerada por computador. Com esse mesmo intuito, Takamatsu *et. al*. (2010) analisaram a variância do ruído de pixels vizinhos, onde foi possível extrair uma característica sobre a relação entre a média dos desvios de ruídos de todos os pixels interpolados e os não interpolados. Kirchner (2010) desenvolveu uma solução aproximada que requer apenas uma operação de filtragem linear em uma imagem.

Em outro aspecto, há métodos que exploram a inconsistência dos artefatos JPEG. Assim, Farid (2009) desenvolveu um método de detecção de adulteração baseado na compressão JPEG. A ideia é que a combinação de duas imagens comprimidas com valores distintos de quantização gera inconsistência de compressão, da mesma forma que há inconsistência no padrão CFA quando há combinação de imagens de modelos de dispositivos diferentes ou clonagem de um objeto em outro ponto da imagem. A detecção de adulteração é dada pela soma do quadrado das diferenças entre cada canal de cor (RGB) e o resultado da compressão utilizando a Transformada Discreta do Cosseno (do inglês, *Discret Cosine Transform,* DCT). Usando uma ideia semelhante, Fridrich *et. al*. (2003) desenvolveram um método de reconhecimento de adulteração do tipo clonagem em imagens JPEG baseada na correlação entre os segmentos da imagem original e o trecho tido como suspeito. Este método verifica a taxa de correlação entre os segmentos da imagem classificando-o como adulterado onde houver baixa correlação. Para isso, aplica-se a Transformada Discreta de Fourier (do inglês, *Discret Fourier Transform*, DFT), nos segmentos. A análise é feita nos picos do sinal gerado que são movidos durante o processo de manipulação da imagem. He *et. al*. (2006) propuseram um método que detecta a adulteração em imagens JPEG utilizando dupla quantização e observando os coeficientes gerados pela DCT. Histogramas com maior periodicidade dos artefatos JPEG indicam uma possível adulteração.

Existem outras formas de autenticar uma imagem, sendo uma delas a utilização de marcações (*watermarking*), porém, poucas câmeras digitais possuem esse recurso, além disso, sua implementação nos sensores causaria perda na qualidade da imagem e custo. Além da autenticação, há outra forma de reconhecer o origem do documento através do cabeçalho EXIF da imagem, mas não é levado em consideração em um cenário forense devido a facilidade de modificar ou destruir tal informação.

A proposta do método aqui apresentada é baseada na técnica descrita por Li *et. al.* (2008) que a, princípio, identifica os blocos da imagem JPEG e em seguida verifica a presença ou ausência de inconsistência dos artefatos. Para o reconhecimento de adulteração, nós identificamos o BAG (*Block Artifact Grid*) do mesmo modo que Li *et. al.* (2008), e localizamos o trecho adulterado marcando os pontos que diferem do bloco JPEG. O BAG da imagem trata-se da demarcação dos blocos JPEG da imagem que são movidos durante a manipulação do tipo clonagem e composição. A Figura 2 mostra um exemplo de inconsistência do BAG.

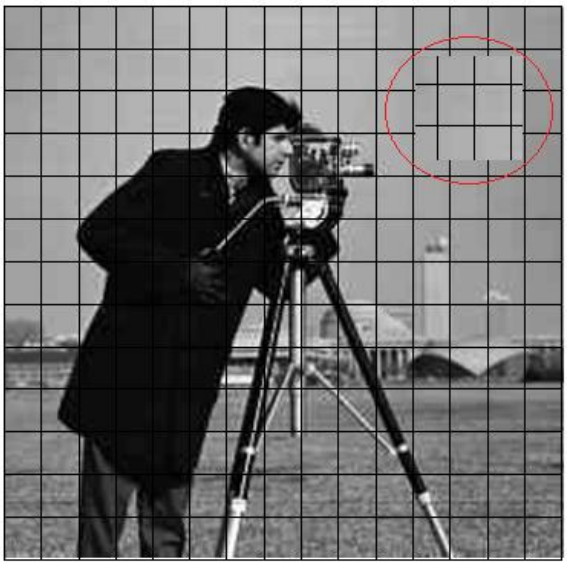

Figura 2. No círculo vermelho pode ser observado um exemplo de inconsistência do BAG (Block Artifcat Grid) gerada a partir da utilização da técnica de manipulação clonagem.

A nossa abordagem é descrita na Seção 3 e os Resultados e Considerações Finais são apresentados nas Seções, 4 e 5, respectivamente.

2. Descrição do método

Ao utilizar uma DCT2 bidimensional em uma imagem ou em blocos, os coeficientes AC depois da quantização tendem a ter valores próximos a zero, enquanto os coeficientes DC tendem a obter valores maiores por conter grande parte das informações presentes na imagem. A partir dessa observação é possível realizar a extração do BAG, pois os valores dos coeficientes AC iguais ou próximo de zero indicam a posição correta do BAG, caso contrário, indicam que há um deslocamento do bloco da imagem. Um caso particular a ser observado são as imagens com alta grau de textura e bordas, pois ao ser aplicado a DCT2 nessas regiões, os valores obtidos estarão localizados nos coeficientes AC após a quantização. Apesar disso, não se torna uma limitação para o método de extração do BAG.

Para iniciar a localização do BAG, deve-se calcular o Efeito Local (do inglês, *Local Efect*, LE) da imagem, como é descrito por Li *et. al*. (2008) de acordo com a seguinte fórmula (1):

$$LE = \sqrt{\frac{\sum_{i=7\ \text{and}\ /\ \text{or}\ j=7} S_{ij}^{\,2}}{S_{00}^{2}}} \qquad \textbf{(1)}$$

O LEi, (para i=1, 2, 3 … n.) é definido como a raiz da soma da última linha e última coluna dividida pelo coeficiente DC de uma janela deslizante 8x8 da imagem em análise. O método percorre a imagem em blocos 8x8 aplicando a DCT2 em cada bloco da imagem, como mostra a fórmula (2). Para cada ponto da imagem é obtido um valor de LE.

$$S_{uv} = \frac{\sqrt{\alpha_u \alpha_v}}{8} \sum_{i=0}^{7} \sum_{j=0}^{7} s_{ij} \cos\frac{u(2i+1)\pi}{16} \cos\frac{v(2j+1)\pi}{16} \qquad \textbf{(2)}$$

Onde:

$$\alpha u = \begin{cases} 2 \text{ se } u = 0 \\ 1 \text{ se } u \neq 0 \end{cases} \qquad \textbf{(3)}$$

$$\alpha v = \begin{cases} 2 \text{ se } v = 0 \\ 1 \text{ se } v \neq 0 \end{cases}$$

O BAG da imagem está localizado nos pontos de valores mínimos que podem ser obtidos utilizando a fórmula (4) para uma imagem de tamanho RxC:

$$S(k,l) = \sum_{i=0}^{\frac{R}{8}-1} \sum_{j=0}^{\frac{C}{8}-1} LE\ (i*8+k, j*8+l) \quad \textbf{(4)}$$

Para k= 0, 1,.., 7 e l= 0, 1,.., 7

$$\mathrm{M}\ (\mathrm{k},\ \mathrm{l}) = \frac{S\ (k,l)}{N} \quad \textbf{(5)}$$

Onde:

- N é o total de blocos 8x8 da imagem.
- M(k,l) é a média dos LE's, para k=0, 1, … 7 e l=0, 1 .. 7.

$$(\mathrm{k}^*,\ \mathrm{l}^*) = \arg\min_{\mathrm{k,l}} \mathrm{M}_{\mathrm{k,l}} \quad \textbf{(6)}$$

Para localizar o BAG, verifica-se qual é a posição de linha e coluna a menor média de LE ocupada na matriz M (k,l), como mostrado na fórmula (6). Os pontos são marcados bloco a bloco com 255 a partir dessas posições encontradas. O reconhecimento da adulteração na imagem é realizado após essa identificação, verificando se os mínimos LE da imagem estão localizados em pontos de BAG. Caso não estejam, são marcados atribuindo o valor de 255 na imagem, indicando que há inconsistência dos artefatos JPEG. A Figura 3 ilustra esse processo.

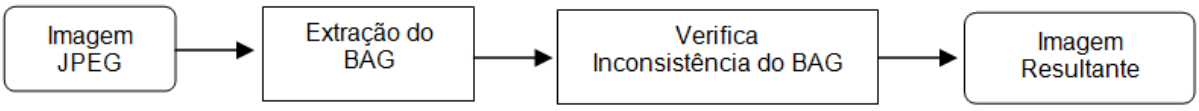

Figura 3. Etapas seguidas para o Reconhecimento de Adulteração em imagens através da inconsistência dos blocos JPEG.

3. Resultados e discussões

Como descrito, a nossa abordagem reconhece uma adulteração em uma imagem JPEG a partir da inconsistência do BAG. Os testes foram realizados em 60 imagens manipuladas utilizando as técnicas de clonagem e composição.

A Figura 4 (a) apresenta a imagem de Lena no formato JPEG com taxa de compressão 30 e na Figura 4 (b) temos o BAG extraído utilizando a fórmula (4). Para extração do BAG foi obtido uma taxa de 98% a 100% de acurácia para as taxas de compressão variando de 10 a 90. Para imagens com compressão 100, a taxa de acurácia foi reduzida para 75%. Isso ocorre devido a compressão ser muito baixa ocasionando pouco ou nenhuma blocagem JPEG. Dessa forma, o método de extração do BAG é mais indicado para imagens com taxa de compressão inferior a 90 em que há maior compressão do sinal após a quantização.

a) **b)**

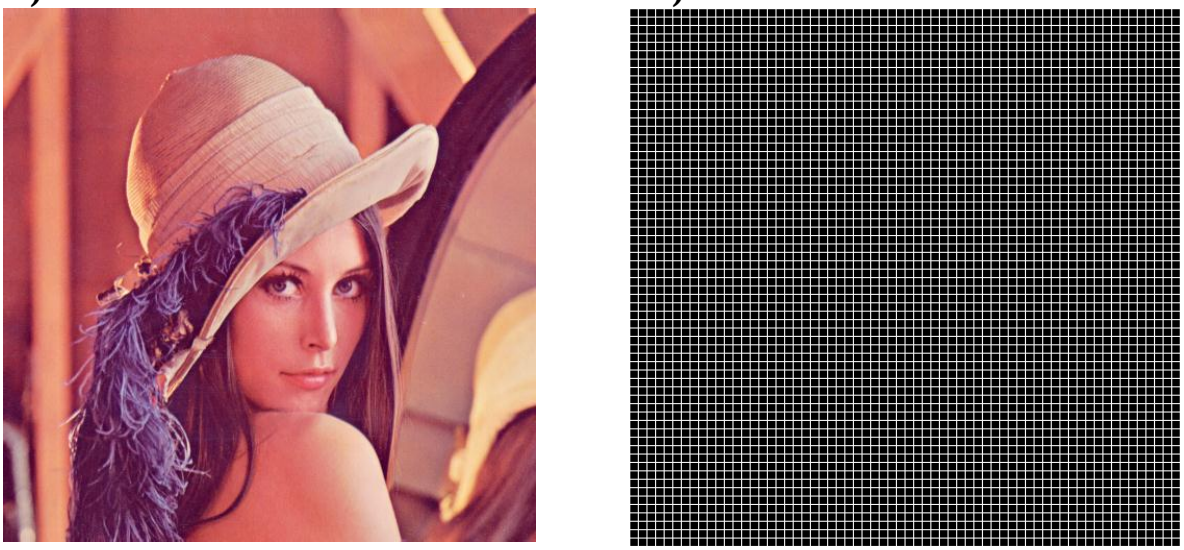

Figura 4. (a) Imagem JPEG Lena. (b) BAG imagem Lena.

A Figura 5 (c) apresenta o resultado de reconhecimento de adulteração para uma imagem manipulada pela técnica de composição. Nesse exemplo, foram combinadas duas imagens iguais, uma com taxa de compressão JPEG 30 e outra com taxa de compressão 100. Como não há perda de informações em imagem com taxa de compressão 100, os coeficientes AC gerados pela DCT2 produzem LE com baixos valores, que, para nosso algoritmo é visto como ponto de BAG.

a) b) c)

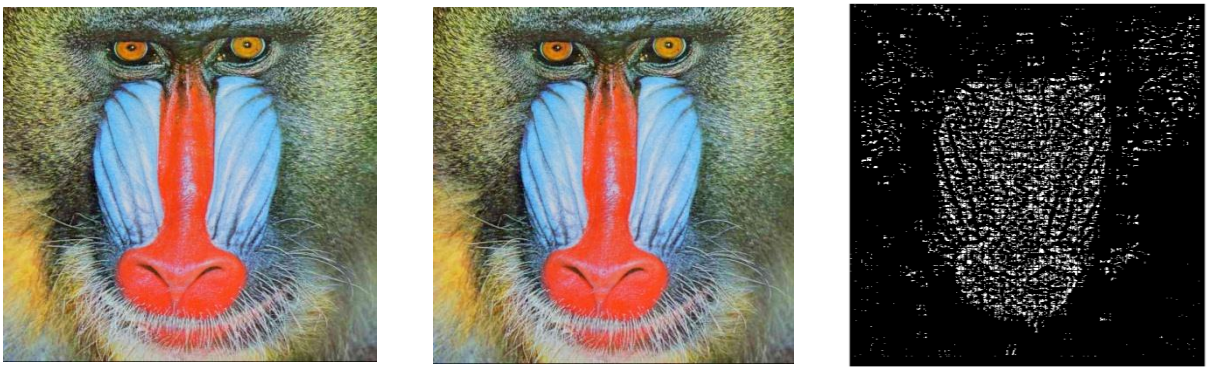

Figura 5. (a) Imagem Macaco Original. (b) Imagem Macaco Manipulada pela técnica de composição. (c) Resultado obtido a partir da verificação da inconsistência dos blocos JPEG.

Assim, como na Figura 5, a Figura 6 (c) apresenta o resultado do reconhecimento de adulteração do tipo composição. Da mesma forma que a imagem anterior, a imagem adulterada é resultado da combinação de duas imagens JPEG com taxas de compressão distintas.

a) b) c)

Figura 6. (a) Imagem original. (b) Imagem manipulada pela técnica de composição. (c) Resultado obtido a partir da verificação da inconsistência dos blocos JPEG.

Na Figura 7 (c) é possível observar o trecho clonado da imagem, além disso, é notável que a adulteração torna-se mais visível quando a imagem é manipulada pela técnica de composição. Apesar das chances serem mínimas, há momentos em que o BAG da imagem pode casar perfeitamente com a área adulterada, principalmente em casos de manipulação do tipo clonagem. Além disso, nossa abordagem limita-se a imagens que possuem muitas regiões de texturas e compressão muito alta, fazendo com que o método encontre vários mínimos e, consequentemente, indicando-os como adulteração.

a) b) c)

Figura 7. (a) Imagem original. (b) Imagem manipulada pela técnica de clonagem. (c) Resultado obtido a partir da verificação da inconsistência dos blocos JPEG.

4. Considerações finais

Com o grande poder de processamento dos softwares de edição, tornou-se mais comum a circulação de imagens manipuladas, e que na maioria das vezes, é feita para modificação do contexto da cena. O reconhecimento de adulterações tem grande importância para o cenário forense, principalmente quando se trata de sua utilização como provas judiciais. Nesse trabalho foi descrita uma abordagem baseada no reconhecimento de adulteração em imagens a partir da análise da inconsistência dos artefatos JPEG, em que, a partir da identificação da marcação do início dos blocos da imagem, foi possível verificar possíveis inconsistências. O método desenvolvido mostrou apenas como limitação imagens com regiões de textura com alto nível de compressão e imagens sem compressão. A partir dos testes realizados foi possível obter taxa de acurácia de 98%, o que o torna competitivo quando comparado com os métodos presentes na Literatura que variam entre 70% a 99%.

Como trabalho futuro, pretende-se utilizar uma abordagem complementar para o reconhecimento de adulterações em imagens sem compressão. De forma semelhante, a análise da inconsistência do padrão CFA é um caminho que pode tornar essa abordagem completa.

5. Agradecimentos

Agradecimento à CAPES pela concessão da bolsa de estudos que permitiu o desenvolvimento desse trabalho.

6. Referências

BAYRAM, S., SENCAR, H. T. e MEMON, N. D. **Classification of digital camera-models based on demosaicing artifacts**. Digital Investigation, 5(1-2):49–59, 2008.

BURVIN, P. Sabeena. ESTHER, J. Monica. **Analysis of Digital Image Splicing Detect**. IOSR Journal of Computer Engineering (IOSR-JCE). e-ISSN: 2278-0661, p- ISSN: 2278-8727 Volume 16. April 2014.

FARID, H. **Exposing Digital Forgeries from JPEG Ghosts**. *IEEE Transactions on Information Forensics and Security*, 4(1):154-160, 2009.

FRIDRICH, Jessica. SOUKAL, David. LUKAS, Jan. **Detection of Copy-Move Forgery in Digital Images**. Proceedings of Digital Forensic Research Workshop. 2003.

GALLAGHER, A. e CHEN, T. **Image authentication by detecting traces of demosaicing**. In Computer Vision and Pattern Recognition Workshops. CVPRW '08. IEEE Computer Society Conference on, pages 1–8, june 2008

HE, J.F., LIN Z.C., WANG L.F., e TANG X.O. **Detecting doctored JPEG images via DCT coefficient analysis**. Lecture Notes in Computer Science, Springer Berlin, vol. 3953, pp.423-435, 2006.

KIRCHNER, Matthias. **Efficient estimation of CFA pattern configuration in digital camera images**. volume 7541, page 754111. SPIE, 2010.

LI, Weihai. YUAN, Y. e YU, Neghai. **Detecting Copy-Paste Forgery of JPEG Image via Block Artifact Grid Extraction**. International Workshop on Local and Non-Local Approximation in Image Processing. 2008.

POPESCU, A. FARID, Hany. **Exposing digital forgeries in color filter array interpolated images**. Signal Processing, IEEE Transactions on, 53(10):3948 – 3959, oct. 2005

ROCHA, Anderson, GOLDENSTEIN S. **CSI: análise forense de documentos digitais**. Belo Horizonte: Sociedade Brasileira de Computação (SBC); 2010. cap. 6, p.263-317: Atualizações em informática.

TAKAMATSU, J., MATSUSHITA, Y. OGASAWARA, T. IKEUCHI, K. **Estimating demosaicing algorithms using image noise variance**. In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pages 279 –286, june 2010.