

# DEEP LOOK IN BLUR ROBUST AND COLOR CONSTANT IMAGE DESCRIPTION

Uma análise detalhada

Álvaro João Silvino (*Autor*)

Centro de Informática-Cin  
UFPE – Universidade Federal de Pernambuco  
Recife, Brasil  
ajsss@cin.ufpe.br

**Abstract**— Image Descriptors is a field growing in many different aspects since for improvement of object recognition accuracy or create indexes for millions of images available in cloud systems. Make color constant descriptors robust to blurs scene becomes more important [9], blurred data is constant in several images, since artificial blur applied by image editor or even in movement of camera or by wrong focus settings, to overcome general image representations.

**Keywords**— *object recognition, blurred imagens, color invariants, color constant descriptor, image representations, .*

## I. INTRODUÇÃO

Imagens originadas de câmeras digitais são dependentes de três fatores principais: o conteúdo físico da imagem (objetos caracterizados na imagem), iluminação (posição e intensidade da fonte luminosa) e características da câmera (configurações espaciais das câmeras durante a captura da imagem). Com a mudança de iluminação um efeito na alteração das cores inevitavelmente ocorre sobre as cores originais dos objetos em cena; Esse problema não é enfrentado pelos seres humanos, onde os olhos possuem uma incrível habilidade de manter as características das cores em diversas formas e formatos de iluminação.

Cores em imagens são informações simples e muito objetivas para caracterizar objetos e seu reconhecimento, esse comportamento já foi comprovado em diversos artigos [1-8]. Da mesma forma que os seres humanos possuem a habilidade de reconhecer as cores em diferentes ambiente iluminados, o computador também precisa dessa capacidade para remover a influência da iluminação em objetos e corretamente indexar ou classificar os objetos.

Existem duas estratégias possíveis para solucionar esse problema: A primeira estratégia consiste em realizar uma direta relação entre os valores aferidos nos objetos e compara-los em relação a um objeto canônico de iluminação, objeto universal de tradução as iluminações. A segunda estratégia consiste em representar imagens através de uma característica criada a partir da imagem, onde é completamente invariante da iluminação e

sua fonte, essa característica é chamada de “color constante descriptor” ou “color invariante”, removendo assim a dependência dos algoritmos de “color constancy” ou consistência de cor [1-8]. Nesse artigos é explorado a segunda estratégia explicada.

J. van de Weijer et al. [6] introduziu uma melhora nas proporções de derivadas de imagem para “edge-based color constant descriptors”. Esse métodos foca na criação de descritores mais robustos com imagens em diferentes efeitos luminosos, diferentes ângulos e principalmente com efeito de embasamento. Representações de descritores baseados em derivadas possuem esse comportamento indesejado que varia com imagens borradas.

Adicionalmente em métodos robustos, o artigo “Blur Robust and Color Constant Image Description”[6] ainda conta com uma proposta em corrigir o efeito de embasamento sobre as bordas entre regiões das imagens usando a ferramenta de “color angles”. A imagem borrada, devido efeitos sintéticos ou por movimento da câmera, faz com que as regiões de bordas das imagens sejam prejudicadas. Assim influenciando de forma mínima as cores, porém irá alterar fortemente a nitidez de transição entre as regiões de transições.

## II. EDGE-BASED COLOR CONSTANCY

Considerando a resposta do sensor  $C \in \{R, G, B\}$  de uma camera com sensibilidade espectral  $f^c$ :

$$C(x) = m^b(x) \int b(\lambda, x) e(\lambda) f^c(\lambda) d\lambda \quad (1)$$

Onde  $m^b$  representa as mudanças geométricas causadas pela geometria da iluminação, superfície geométrica iluminada.  $b$  é a superfície de albedo,  $e(\lambda)$  é a distribuição espectral da iluminação. Para simplificar e aproximar essa função levaremos os seguinte equação [1,2,3,4]:

$$C(x) = m^b(x) b^c e^c \quad (2)$$

Onde,  $b^C(x) = b(\lambda^C, x)$ , and  $e^C = e(\lambda^C)$ ,  $b^C(x)$  representa os valores de albedo na distribuição espacial da imagem em cada distribuição espectral de canal disponível  $C$ .

Inicialmente, consideramos que a iluminação geométrica, ou a geometria da cena iluminada é sempre constante, dessa forma  $m^b$  deve ser invariante e independente de  $x$ , essa consição é chamada de Mondrian ou flat world assumption [1]. Funt e Finlayson [2] mostraram que com essa condição a derivada do logaritmo da resposta de um sensor de câmera é invariante e independentemente da cor iluminada da cena, assim:

$$\frac{\partial}{\partial x} \ln C(x) = \frac{\partial}{\partial x} (\ln b^C(x) + \ln m^b(x) + \ln e^C) = \frac{b_x^C}{b^C} \quad (3)$$

Sendo  $b_x^C$  o valor de albedo para cada canal  $C$  e o subscript  $x$  representa a variação espacial na imagem e  $b^C$ , representa uma constante presente em cada canal  $C$ . Assim a representação dos canais de cores são:  $p = \{p_1, p_2, p_3\} = \{\frac{R_x}{R}, \frac{G_x}{G}, \frac{B_x}{B}\}$ .

Para Gevers e Smeulders [3,4] não é considerado as condições Mondrian, e sim é considerado o caso 3D world [3]. Onde  $m^b(x)$  varia em relação a  $x$ , quando a orientação do objeto é mudada essas condições são levadas em conta. Dessa forma os diferentes planos de orientação irão sofrer variações no termo  $m^b$ .

Para garantir maior robustez e invariância, Gevers e Smeulders propuseram o seguinte equação[4]:

$$\frac{\partial}{\partial x} \ln \frac{C(x)}{D(x)} = \frac{b_x^C}{b^C} - \frac{b_x^D}{b^D} = \frac{b_x^C b^D - b^C b_x^D}{b^C b^D} \quad (4)$$

Onde  $D \in \{R, G, B\}$  e  $D \neq C$ . Para os três canais de cores consequentemente levam para duas razões independentes  $m = \{m_1, m_2\} = \{\frac{R_x G - G_x R}{RG}, \frac{G_x B - B_x G}{GB}\}$ .

### III. BLUR RPBUST AND COLOR CONSTANT IMAGE DESCRIPTION

Embasamento de imagens pode ter diversas formas de ocorrer. Desde movimento da cena ou da câmera durante o momento de captura da imagem[6].

Todos os métodos acima discutidos tem como principal ferramenta a aplicação de derivadas. O principal objetivo de uma derivada em uma imagem é a clara definição das bordas (comumente chamados de “edges”). O embasamento ataca não só a clareza e nitidez da imagem como também degrada as regiões de mudança ou as bordas da cena da imagem.

Para conseguir robustez em relação a embasamento, foi proposto a seguinte equação:

$$\varphi_p^1 = \arctan\left(\frac{p_1}{p_2}\right), \varphi_p^2 = \arctan\left(\frac{p_2}{p_3}\right) \quad (5)$$

Onde gera o seguinte angulo de cor  $\varphi_p = \{\varphi_p^1, \varphi_p^2\}$ .

Uma outra estratégia similar para gerar constância de cor para imagens embasadas é a utilização das razões invariantes a geometria de iluminação da cena como mostrado no **Eq.4**. Dessa forma foi proposto a seguinte equação:

$$\varphi_m = \arctan\left(\frac{m_1}{m_2}\right) \quad (6)$$

Quando é utilizado as cores angulares ou ângulo de cor, proposto nas **Eq. 5** e **Eq. 6** a dependência do embasamento dentro da imagem é diminuída e atenuada.

### IV. METODOLOGIA

Nesse artigo foi aplicado os quatro descritores de imagem em um simples classificador K-NN (“K – Nearest Neighbor”) [5,7,8] com objetivo de testare medir se o processo de indexar as imagens corretamente.

Para a construção dos descritores em todos os experimentos foram aplicados a função de logaritmo neperiano com a soma de uma constante 0.001 para evitar logaritmos de zero que tentem ao infinito. Além, para o componente de derivada foi utilizado a derivada laplaciana de segundo grau com desvio padrão  $\sigma = 2$  durante a aplicação de derivada em todos os métodos descritos (**Eq.3**, **Eq.4**, **Eq.5**, **Eq.6**).

Por fim esses dados foram normalizados e construído um histograma N-dimensional que representa o descritor final da imagem, para o método  $p = \{p_1, p_2, p_3\}$  foi utilizado um histograma com 3 dimensões,  $m = \{m_1, m_2\}$  e  $\varphi_p = \{\varphi_p^1, \varphi_p^2\}$  com 2 dimensões e por fim  $\varphi_m$  com 1 dimensão. Para comparar cada descritor foi utilizado a distância euclidiana[5,6]. A métrica utilizada foi a aplicação de um classificador simples K-NN onde o K pode varia de acordo com os experimentos. K=1 necessariamente existe apenas 1 vizinho correspondentes a classe do objeto teste, ou seja para cada imagem ou objeto teste existe apenas uma imagem a ser encontrada. Para  $K > 1$  deve existir mais de 1 objeto correspondente a classe de teste. Os resultados apresentados são a taxa de acerto para cada método e seus K vizinhos.

### V. RESULTADOS

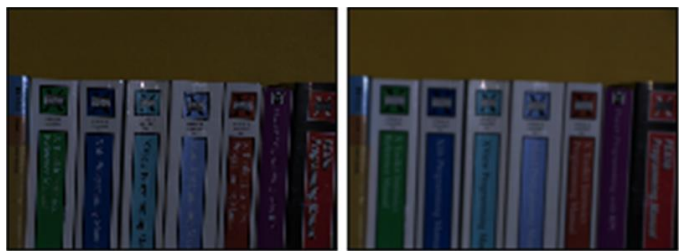
Foram utilizadas 2 bases durante os experimentos, 1 das bases utilizadas é da instituição Saimon Fraser[10] e dedicadas em variação de iluminação em múltiplos objetos e a última foi retirada do estudos de Joost van de Weijer[11].



**Figura 1:** Pequena representação dos objetos das bases de imagens. Primeira linha corresponde ao objetos de estudo de Weijer [11] chamada Real World Blurred Data. Segunda linha

são os objetos da base de dados da Saimon Fraser[10] essas bases são compostas por diversos objetos sobre diversas iluminações diferentes.

Primeiro experimento foi dedicado a testar a robustez de cada descritor em relação a embasamento artificial. Dessa forma extraímos 20 objetos sobre a mesma variação de iluminação e assim aplicamos um filtro gaussiano com desvio padrão  $\sigma = 2$  para embasar artificialmente as imagens totalizando 40 imagens [10]. O objetivo do experimentos é aplicar os 4 métodos e faze-los achar suas contrapartes correspondentes.



**Figura 2:** Exemplo de uma imagem retirada da base[10] e sua contraparte embasada.

Método	K=1
$p$	0.150
$m$	0.050
$\varphi_p$	0.800
$\varphi_m$	0.750

**Tabela 1:** Aplicação de um classificador 1-NN, é possível ver a taxa de acerto e a robustez relacionada aos ângulos de cor já se destacam.

No experimento seguinte utilizamos uma base chamada “Fluorecente iluminance variance”, disponível no repositório da Simon Fraser [12]. Nessa base possui 6 objetos sobre a influência de 9 variações de iluminação diferentes.

Método	K=1	K=3	K=5
$p$	0.833	0.667	0.500
$m$	0.667	0.333	0.250
$\varphi_p$	1.000	1.000	1.000
$\varphi_m$	1.000	1.000	1.000

**Tabela 2:** Aplicamos o K-NN para  $k=[1,3,5]$ . Para os métodos de cores angulares o resultado foi a correta classificação de todos os objetos, isso é, o método conseguiu indexar todos os objetos corretamente.

Para se verificar a robustez em relação a rotação e escala dos descritores de imagens, foi realizado um experimento semelhante ao Tabela 1, onde foram selecionados 20 objetos da base Saimon Fraser [10] e é aplicado uma transformação afim

em cada uma das imagens, experimento semelhante em artigos de Xiong e Xu [7] como a Figura 3 sugere.



**Figura 3:** Um exemplo de uma imagem com a transformação afim [7,8].

Método	K=1
$p$	0.200
$m$	0.050
$\varphi_p$	0.400
$\varphi_m$	0.150

**Tabela 3:** Resultado da classificação com transformação afim.

Apesar dos resultados inferiores aos experimentos anteriores sugerem que os descritores de cores angulares ainda são superiores aos métodos prévios, porémé verificado que ainda há muito espaço para melhorias, como o estudo de Xiong e Xu [7] e Xu e Li [8] exploram em um de seus estudos.

Adicionalmente aos testes de embasamento, a base chamada “Real world Blured Data” [11] é focada em imagens com e sem embasamento, um conjunto de 20 objetos totalizando 40 imagens. O diferencial dessa base é que o embasamento fora provocado pela própria câmera ou alguma condição do movimento no momento da captura da imagem. O experimento segue igual ao do embasamento pela transformação gaussiana na Tabela 1. Uma imagem de teste a procura de sua contraparte única.

Método	K=1
$p$	0.200
$m$	0.200
$\varphi_p$	0.850
$\varphi_m$	0.550

**Tabela 4:** Resultado do classificador 1-NN para a base Real world Blured Data de Weijer [11].

Como mostrado na Tabela 4 os métodos angulares são superiores nessas condições, em especial o método  $\varphi_p$  [6].

Com o objetivo de testar a robustez em grande massa de imagens o próximo experimento foi organizado. Unindo diversas bases disponíveis em Simon Fraser [10] foi reunido 56 objetos diferentes sobre a influência de 11 fontes de iluminações

diferentes totalizando 616 imagens diferentes, é importante informar que as 11 fontes de iluminações são consistentes e iguais aos 56 objetos reunidos.

Para esse experimento final foi utilizado o classificador K-NN para  $k=[1,3,5]$ .

Método	K=1	K=3	K=5
$p$	0.393	0.125	0.054
$m$	0.214	0.107	0.036
$\varphi_p$	0.821	0.661	0.607
$\varphi_m$	0.643	0.393	0.339

**Tabela 5:** Resultado do experimento reunindo 56 objetos sob influência de 11 iluminações diferentes.

## VI. CONCLUSÃO

Nesse artigo os ângulos de cores são propostos como descritores invariantes a iluminação e robustos a imagens embasadas ou degradadas através do embasamento, de forma artificial com utilização de filtros ou por embasamentos naturais com movimento de câmera. É importante verificar que esse método de invariância de iluminação não necessita dos parâmetros da câmera ou da iluminação da cena.

Através dos resultados obtidos a utilização dos ângulos de cor  $\varphi_p$  e  $\varphi_m$  [6] em todos os experimentos demonstrou construir descritores mais poderosos e expressivos, auxiliando um simples classificador em ter uma taxa de acerto muito superior que os métodos anteriores.

## VII. ESTUDOS FUTUROS

Os estudos futuros podem envolver extratores de características com métodos computacionais mais sofisticados em definir as fronteiras[8], ou podemos explorar outros extratores de fronteiras diferentes de derivada, como operadores morfológicos.

## REFERÊNCIAS

- [1] Swain, M. J.—Ballard, D. H.: Color Indexing. International Journal of Computer Vision, Vol. 7, 1991, No. 1, pp. 11–32.
- [2] Funt, B. V.—Finlayson, G. D.: Color Constant Color Indexing. IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol. 17, 1995, No. 5, pp. 522–529.
- [3] Gevers, T.—Smeulders, A.: Color Based Object Recognition. Pattern Recognition, Vol. 32, 1999, pp. 453–464.
- [4] Th. Gevers and A. Smeulders, “Color based object recognition,” Pattern Recognition, vol. 32, pp. 453–464, Mar. 1999.
- [5] Adjero, D. A.—Lee, M. C.: On Ratio-Based Color Indexing. IEEE Trans. On Imaging Processing, Vol. 10, 2001, No. 10, pp. 36–48.
- [6] Van de Weijer, J.—Schmid, C.: Blur Robust and Color Constant Image Description. Proc. Int. Conf. on Image Processing, Oct. 2006, Atlanta, pp. 993–996.
- [7] Li, B.—Xiong, W.—Xu, D.: A Supervised Combination Strategy for Illumination Chromaticity Estimation. ACM Trans. on Applied Perception. Vol. 8, 2010, No. 1.
- [8] Bao, H — Xu, D. — Li, B.: Color Constant Descriptors Combining Image Derivative Structures Proc. Computing and Informatics, Vol. 31, 2012, 971–982.
- [9] B. Yoshua, C. Aaron and V. Pascal, Representation Learning: A Review and New Perspectives, Apr. 2014
- [10] Simon Fraser University; Computational Colour Science. Available from: [http://www.cs.sfu.ca/~colour/data/objects\\_under\\_different\\_lights](http://www.cs.sfu.ca/~colour/data/objects_under_different_lights)
- [11] Van de Weijer, J. (2006, Oct). Real World Blurred Data. Available from: [http://lear.inrialpes.fr/people/vandeweyer/blur\\_data/blur.html](http://lear.inrialpes.fr/people/vandeweyer/blur_data/blur.html)
- [12] Simon Fraser University; Computational Colour Science. Fluorescent Surfaces data Available from: [http://www.cs.sfu.ca/~colour/data/colour\\_constancy\\_test\\_images/fluorescent/](http://www.cs.sfu.ca/~colour/data/colour_constancy_test_images/fluorescent/)