

**Uma métrica para avaliação
de diferença perceptual em imagens
para deficientes visuais de cores**

Rafael de Oliveira Lopes Gonçalves

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Ciência da Computação
Orientador: Prof. Dr. Roberto Hirata Jr.

São Paulo, Agosto de 2013

**Uma métrica para avaliação
de diferença perceptual em imagens
para deficientes visuais de cores**

Esta é a versão original da dissertação elaborada pelo
candidato Rafael de Oliveira Lopes Gonçalves, tal como
submetida à Comissão Julgadora.

RESUMO

As diversas técnicas de recoloração para a melhoria de visualização imagens para o indivíduo deficiente visual de cores podem permitir o entendimento de uma informação contida em uma imagem. No entanto, alteram a percepção da imagem. Esta diferença perceptual nem sempre é desejada ou mesmo necessária. Neste trabalho é proposta uma métrica ao problema: avaliar a diferença perceptual da imagem e assim possibilitar a identificação automática quais imagens apresentam problemas para visualização pelo deficiente visual de cores e quando é necessária a aplicação de uma técnica de recoloração. Neste trabalho desenvolvemos uma utilizando um modelo de atenção visual para mensurar a diferença perceptual entre a visualização de um indivíduo e um deficiente para visão de cores. Esta métrica foi testada verificando-se que seu valor diminui com o uso de técnicas de recoloração previamente validadas. Ao utilizar um modelo computacional de saliência, este trabalho prediz uma diferença de atenção exógena entre indivíduos normais e indivíduos com deficiência na visão de cores.

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Descrição do problema a ser tratado	1
1.2	Trabalhos relacionados	6
1.3	Contribuição do trabalho	6
1.4	Organização do texto	6
2	VISÃO A CORES - HUMANA E COMPUTACIONAL	8
2.1	Princípios da visão humana de cores	8
2.1.1	Visão Anômala de Cores	10
2.1.2	Deficiência de visão verde-vermelho	10
2.1.3	Deficiência de visão azul-amarela	11
2.1.4	Acromatopsia e Monocromacia	11
2.2	Simulação de Visão Deficiente	14
2.3	Recoloração de imagem para visualização deficiente	16
3	ATENÇÃO VISUAL E SALIÊNCIA	19
3.1	Atenção Visual	19
3.2	Modelos de Saliência	20
4	MODELO DE SALIÊNCIA BASEADO EM DICIONÁRIO	24
4.1	Framework de saliências baseadas em dicionário de imagens	24
4.1.1	Análise de componentes Independentes	25
5	APLICAÇÃO DO MODELO DE SALIÊNCIA PARA CÁLCULO DA DIFERENÇA DE VISAO NORMAL E DEFICIENTE	29
5.1	Modelo de visão deficiente	29
5.2	Calculando a diferença perceptual entre a visão normal e deficiente	30
5.2.1	Experimento	31

6 CONCLUSÃO 41

REFERÊNCIAS BIBLIOGRÁFICAS 43

LISTA DE FIGURAS

Figura 1	Reprodução de imagem de jogo Call of Duty onde o mapa no canto superior esquerdo apresenta problemas de acessibilidade (Adaptado de http://jogos.uol.com.br/).	2
Figura 2	(Fonte [18]).	4
Figura 3	(Fonte [18])	5
Figura 4	Curvas de absorção normalizadas dos três tipos de pigmentos de cores. (Adaptado de Bowmaker [3]).	9
Figura 5	Curvas de absorção espectral relativa de fotorreceptores de portadores de deficiência visual de cores. No quadro A as curvas de absorção típicas dos cones de um indivíduo com protanopia sem a presença do cone L. Da mesma forma está ausente a curva do cone M no quadro B de deuteranopia. Os quadros C e D apresentam as curvas de absorção com curvas anômalas L e M respectivamente. (Fonte:[9])	12
Figura 6	Simulações de diversos níveis de deuteranomalias e protanomalias (Fonte:[27])	13
Figura 7	Simulações de visão deficiente utilizando o método de Brettel et al. [4]	15
Figura 8	Sensibilidade espectral nos eixos de oponência (Fonte:[27])	16
Figura 9	Imagem recolorida utilizando o método de [26]	17
Figura 10	Movimentos do olhar de uma pessoa por 2 minutos. (Fonte:[42])	20
Figura 11	(Fonte:[42])	21
Figura 12	Garotinha de vermelho (Fonte:A Lista de Shindler (1993))	22

Figura 13	Mapa de saliência calculado com componentes independentes	28
Figura 14	Comparando mapas padronizados de saliência.	33
Figura 15	Diferenças perceptuais para visão normal e protranope.	34
Figura 16	Diferenças perceptuais para visão normal e deuteranope.	35
Figura 17	Diferenças de mapas normalizados para visão normal protranope e deuteranope.	36
Figura 18	Comparação entre a a diferença dos mapas de atenção e a diferença euclideana.	40

LISTA DE TABELAS

Tabela 1	Imagens filtradas pela técnica de [26] e os respectivos mapas de diferenças.	37
Tabela 2	Imagens filtradas pela técnica de [23] e os respectivos mapas de diferenças.	38

ACRÔNIMOS

DVC	Deficiente Visual de Cores
TN	Tricromata Normal
ICA	Análise de Componentes Independentes

INTRODUÇÃO

Aproximadamente 200 milhões de pessoas no mundo possuem algum tipo de deficiência para a visão de cores[36]. A incapacidade de distinguir as cores de uma imagem pode prejudicar o entendimento da mesma. O indivíduo deficiente pode ter dificuldade para compreender a legenda de um mapa[25] ou outros códigos de cores artificiais[6, 1], pode sofrer dificuldade na vida profissional[7, 2] ou mesmo jogando jogos eletrônicos, como mostra a Fig.1. Na imagem da Fig.1a, o mapa, no canto superior esquerdo da tela, guia o jogador para que este reconheça aliados e inimigos por um código de cor verde/vermelho. A imagem da Fig.1b mostra a simulação da visão da mesma imagem por um deficiente para cores, evidenciando a dificuldade de reconhecer as duas classes de símbolos no mapa.

Do ponto de vista de quem produz o conteúdo visual, seria interessante saber se uma parcela da população sentiria dificuldade para o entendimento da informação.Embora a W3C estipule orientações para guiar a construção de sites acessíveis a deficientes[5] e existam técnicas para a simulação de deficiência visual de cores[4, 27], não existe mecanismo automático para a detecção de problemas de acessibilidade para cores.

Para o deficiente visual, existem diversas técnicas para melhoria de visualização de imagens[18, 14, 23]. Elas permitem que o deficiente tenha acesso à informação contida na imagem antes inacessível. No entanto, muitas vezes o deficiente não sabe quando a aplicação de uma dessas técnicas seria necessária.

1.1 DESCRIÇÃO DO PROBLEMA A SER TRATADO

Neste trabalho propomos um problema a ser resolvido e uma abordagem possível para tratá-lo.



(a) Original



(b) Visão deficiente simulada

Figura 1: Reprodução de imagem de jogo Call of Duty onde o mapa no canto superior esquerdo apresenta problemas de acessibilidade (Adaptado de <http://jogos.uol.com.br/>).

A exemplo do que acontece nos mapas da Fig.1, um indivíduo com deficiência visual para cores ao observar uma imagem pode deixar de compreender alguma informação relevante. Um possível exemplo é a cena escolar de um indivíduo com deficiência para enxergar cores vermelhas e verdes, frente a uma quadro verde não ser capaz de observar as marcas de um giz vermelho. Ou como a cena da Fig.2 onde o indivíduo deficiente enfrentaria dificuldades para perceber as diferentes cores dos lápis. Uma ideia de como um indivíduo poderia perceber a imagem é dada pela simulação da Fig.2b

Uma possível abordagem seria a aplicação de algum filtro existente na literatura de melhoria de visualização para este tipo de deficiência. As imagens das Fig.3, mostram a cena anterior com a aplicação de um filtro de melhoria proposto em [18]. A transformação permitiria a contagem do número de diferentes lápis.

Estes filtros alteram a percepção final da imagem tanto para o deficiente quanto para o indivíduo sem deficiência de maneira que seja possível para o deficiente realizar alguma tarefa que originalmente não era possível. No entanto como as cores na imagem filtrada foram trocadas algumas informações foram perdidas como a sensação intrínseca e subjetiva da cor original no contexto e sua identificação. No exemplo acima, o lápis que criava a confusão de cores para o deficiente, após a aplicação do filtro não poderá mais ter sua cor original identificada por um indivíduo de visão normal e nem causará sensação que a cor original causaria no observador (deficiente ou não). Esta alteração pode ser indesejável e portanto a utilização indiscriminada de um filtro de melhoria deveria ser evitada sendo utilizadas em casos de real necessidade.

Neste trabalho criamos um método para permitir a identificação de possíveis perdas de pontos importantes nas imagens. De posse destas medidas, é possível avaliar a necessidade ou não da aplicação de um filtro de melhoria. E ainda após a aplicação deste filtro de melhoria.



(a) Original



(b) Visão deficiente simulada

Figura 2: (Fonte [18]).



(a) Imagem filtrada



(b) Visão deficiente simulada da imagem filtrada

Figura 3: (Fonte [18])

1.2 TRABALHOS RELACIONADOS

Até onde sabemos, o problema de calcular uma medida que identifique possíveis problemas de visualização para deficientes não é tratado diretamente pela atual literatura. No entanto, diversos trabalhos abordam o tema paralelamente ao tema central. Os trabalhos de recoloração de imagens (melhor discutidos na Sec.2.3) quando automáticos utilizam, explícita ou implicitamente, uma medida de erro a qual em geral é minimizada.

Algumas dessas medidas são definidas para cada cor ou para cada pixel. Diversos autores [24, 23, 18] utilizam a diferença de cada cor da imagem para o indivíduo com visão normal e deficiente. Neste tipo de comparação, as cores são usualmente comparadas pela seu brilho: cores que com pouca diferença de brilho poderiam causar confusão para a visualização de um indivíduo deficiente. Em geral, esta medida não leva em conta as interações entre diferentes regiões de cores e texturas.

1.3 CONTRIBUIÇÃO DO TRABALHO

As principais contribuições deste trabalho incluem:

- Criamos um *framework* para modelos computacionais para a atenção humana baseado em dicionários esparsos. A partir deste *framework* definimos um modelo baseado em análise de componentes independentes.
- Apresentamos uma possível abordagem ao problema de detectar problemas de visualização para deficientes visuais para cores.
- Os resultados apresentados, segundo nosso modelo, exibem diferença de atenção visual entre visão deficiente e normal.

1.4 ORGANIZAÇÃO DO TEXTO

O texto a seguir está organizado da seguinte maneira. No Cap.2 são apresentados conceitos básicos sobre visão de cores e visão deficiente de cores.

Também são apresentados os principais métodos de simulação de visão deficiente e de melhoria de imagem para deficientes. No Cap.3 são discutidos modelos de atenção visual justificando o seu uso no Cap.??, onde é apresentado um *framework* de modelos de atenção visual. Este modelo é aplicado para encontrar diferenças entre a percepção de indivíduos com diferentes sistemas de visão no Cap.5. Neste capítulo são apresentados os resultados dos experimentos realizados. Por fim, no Cap.6 é feita um sumário dos resultados obtidos neste trabalho.

VISÃO A CORES - HUMANA E COMPUTACIONAL

2.1 PRINCÍPIOS DA VISÃO HUMANA DE CORES

A visão de cores não se dá apenas por uma característica da distribuição espectral da luz que alcança os olhos mas também depende de todo o aparato fisiológico humano para criar sua percepção. A luz que chega à retina encontra uma camada densa de células fotorreceptoras chamadas cones.

A absorção ou não da luz pela célula cone e conseqüentemente a criação de potencial elétrico na membrana da célula é determinada pelo seu pigmento de cor[39]. Estes pigmentos ficam localizados no segmento externo da célula e são formados por dois componentes proteicos: 11-cis-retinal e uma opsina.

Num sistema de visão humana normal, há três tipos de fotorreceptores para a visão ¹ de cores os quais são denominados L, M e S (long, medium e short) de acordo com o comprimento de onda onde há o pico de absorção de luz à classe de pigmento correspondente. A Fig.4 mostra as curvas espectrais médias de absorção normalizadas do segmento externo das três classes cones de um indivíduo com visão normal para cores. As curvas com picos de absorção em 420, 534 e 564 *nm* correspondem aos pigmentos do tipo S, M e L respectivamente[37]. Nos seres humanos as opsinas dos pigmentos L e M são 98% similares em suas cadeias de aminoácidos enquanto cada uma é 43% semelhante à opsina do pigmento S[29].

Pelo princípio da univariância, uma vez absorvido, cada fóton tem efeito na célula que independe de seu comprimento de onda[37]. Para cada célula o número de fótons absorvidos pode ser alterado por uma mudança de intensidade de luz ou de comprimentos de ondas do espectro. Sendo as-

¹ Há ainda os bastonetes, contudo estes não têm grande contribuição para a visão de cores.

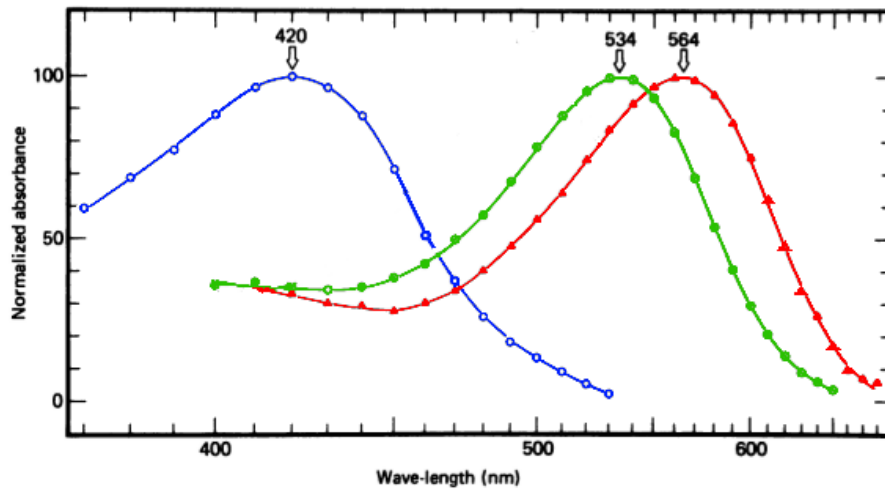


Figura 4: Curvas de absorção normalizadas dos três tipos de pigmentos de cores. (Adaptado de Bowmaker [3]).

sim, cada classe de cone sozinha não pode diferenciar mudanças de cores de mudanças de intensidades, sendo necessária a comparação dos sinais de fotorreceptores de mais de uma classe para criar a visão de cores.

Para diferenciar as cores o sistema visual possui células que comparam os sinais de mais de um tipo de fotorreceptor. Desta maneira, a visão de cores verde-vermelho é formada a partir da comparação entre os sinais dos cones L e M, enquanto a visão de amarelo-azul é dada a partir da comparação entre os sinais dos cones dos cones S e uma combinação de L e M. Ainda assim é possível criar percepções de cores iguais para distribuições espectrais diferentes. Este tipo de fenômeno é chamado de *matamerismo*[41].

O motivo dos bastonetes contribuírem pouco para a visão de cores é justamente por não poderem funcionar ao mesmo tempo que outro fotorreceptor. Os bastonetes funcionam bem em luz escotópica (luz da noite) enquanto os cones funcionam bem em luz fotópica (luz do dia) sendo bem restrita a faixa de intensidade de luz que os cones e bastonetes estão funcionando em conjunto[33].

2.1.1 *Visão Anômala de Cores*

Um ser humano com visão normal para cores pode distinguir por volta de 10 milhões de diferentes cores [21]. Este é chamado de *Tricromata Normal (TN)*, uma vez que possui os três tipos de células cone, seus fotopigmentos não apresentam anomalias e assim pode perceber corretamente todas as cores a partir dos três estímulos fundamentais. A incapacidade de distinguir algumas dessas cores está associada a deficiências genéticas herdadas ou deficiências adquiridas. As deficiências herdadas são normalmente causadas por alterações dos genes que codificam as moléculas dos fotopigmentos dos cones.

Um Deficiente Visual de Cores (DVC) que apresenta somente dois tipos de cones funcionais é chamado de *dicromata*. Há também tricromatas possuidores de anomalias, chamados de *tricromatas anômalos*. Neste caso, a percepção de um dos três estímulos é anômala porém ainda presente. Eles podem criar correspondência de qualquer cor usando os três estímulos mas respondem a estes de maneira diferente da normal. A ausência ou presença anômala de um dos tipos de pigmentos faz com que cores que possuem distribuições espectrais diferentes e percepções diferentes para indivíduos normais tornem-se metamerismos para o DVC.

2.1.2 *Deficiência de visão verde-vermelho*

Por volta de 8% dos homens caucasianos e 1% das mulheres são portadores de algum tipo de anomalia para visão de cores verde/vermelho [32]. Em grande parte, são anomalias congênitas que afetam a codificação das opsinas L e M. O fato dos genes que codificam tais ficarem no cromossomo X e estas anomalias serem recessivas, justificam a grande diferença entre a quantidade de portadores femininos e de portadores masculinos.

Essas anomalias que causam a inexistência de cones do tipo L e M funcionais são chamadas de *protanopia* e a *deuteranopia* respectivamente. Existem ainda indivíduos que possuem uma dessas classes de cones anômala, ou seja o indivíduo possui três tipos de cones no entanto os cones de uma

classe absorvem a luz diferentemente de um tricromata normal, tendo sua curva de absorção espectral deslocada²[12].

Os quadros da Fig.5 mostram as curvas de absorção dos cones e os respectivos deslocamentos causados pelas opsinas alteradas para as diferentes deficiências. O quadro A mostra um possível indivíduo com protanopia, sem cones L. O quadro B mostra indivíduo com deuteranopia, sem cones M. O quadro C mostra indivíduo com protanomalia onde os cones L' estão com a curva de absorção deslocada para menores comprimentos de onda menores, mais próximos de M. Da mesma forma, o quadro D mostra indivíduo com deuteranomalia com a curva de absorção M' deslocada.

A Fig.6 mostra o resultado da simulação de diversos níveis de protanomalia e deuteranomalia conforme o deslocamento de absorção dos cones feita por Machado et al. [27].

2.1.3 Deficiência de visão azul-amarela

A tritanopia é causada por mutações no gene do pigmento S afetando menos do que 1 indivíduo a cada 10000 da população humana[28]. Estes indivíduos possuem dificuldade para diferenciar amarelo de azul. Ao contrário do caso para verde-vermelho é desconhecida a presença de uma forma tricromata anômala desta deficiência.

2.1.4 Acromatopsia e Monocromacia

A acromatopsia herdada é uma deficiência autossômica recessiva rara caracterizada pela ausência de cones funcionais e uma capacidade visual fotópica reduzida. A monocromacia herdada também é uma desordem rara caracterizada pela ausência de cones L e M funcionais podendo ser causada por alterações no cromossomo X.

² O deslocamento da curva de absorção costuma ser referenciada em quantos *nm* de comprimento de onda dista o pico deslocado do pico de um tricromata normal.

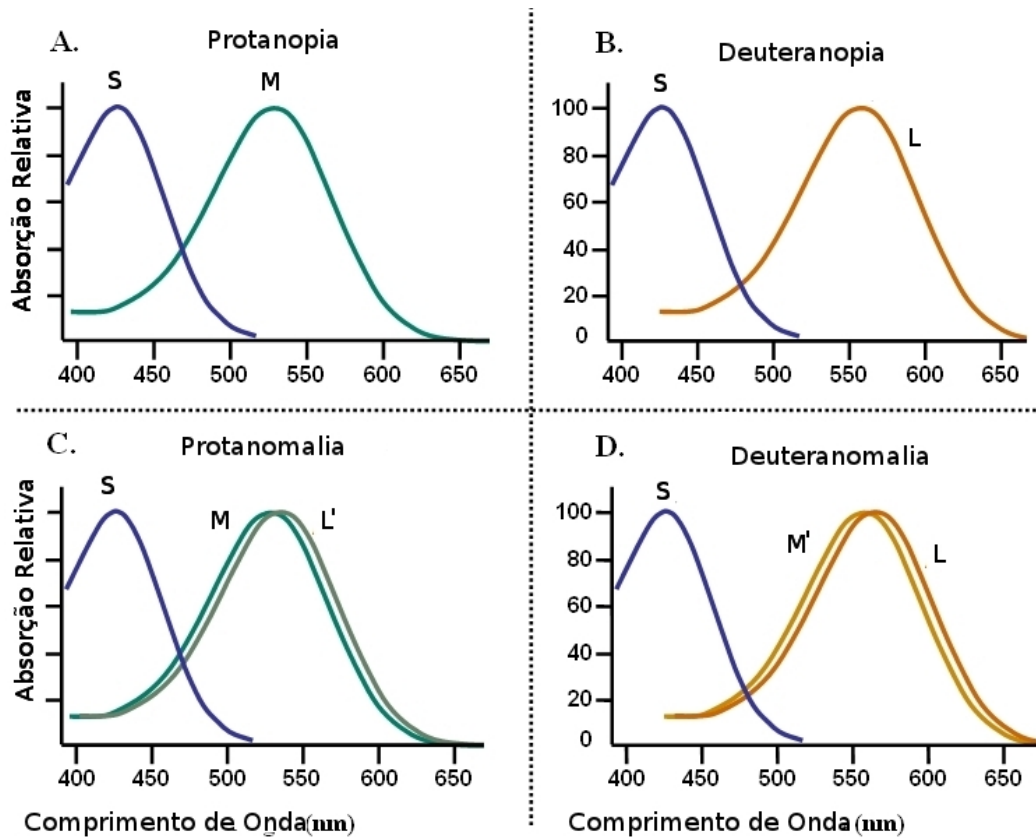
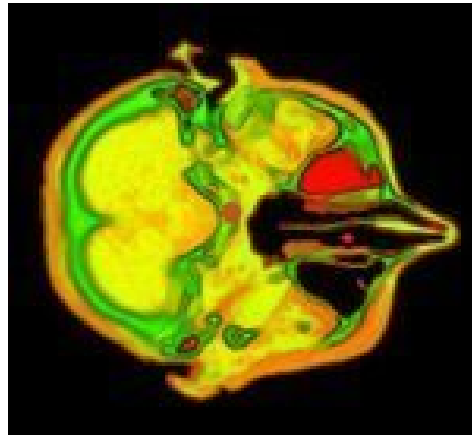
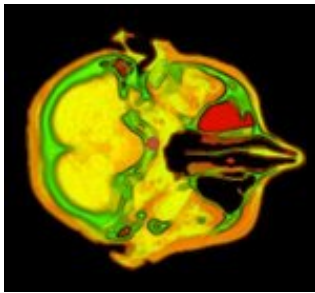


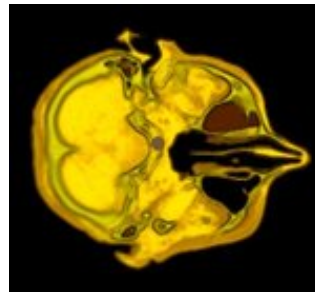
Figura 5: Curvas de absorção espectral relativa de fotorreceptores de portadores de deficiência visual de cores. No quadro A as curvas de absorção típicas dos cones de um indivíduo com protanopia sem a presença do cone L. Da mesma forma está ausente a curva do cone M no quadro B de deuteranopia. Os quadros C e D apresentam as curvas de absorção com curvas anômalas L e M respectivamente. (Fonte:[9])



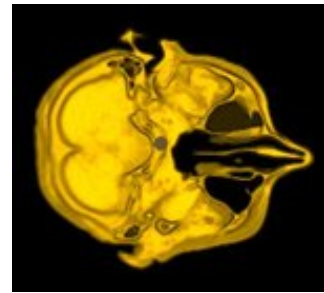
(a) Imagem Original



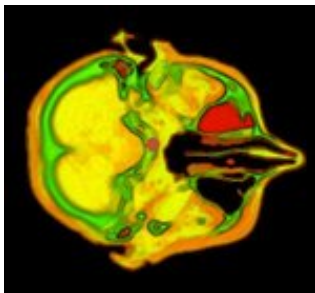
(b) Protanomalia 2nm



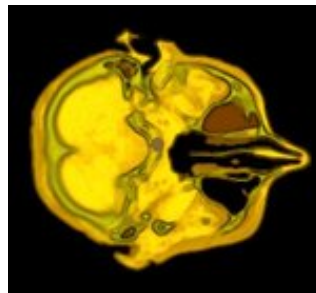
(c) Protanomalia 14nm



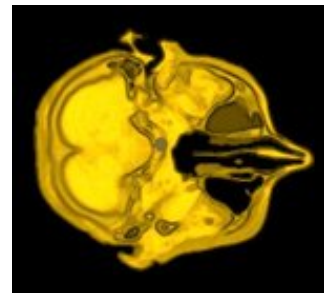
(d) Protanopia



(e) Deuteranomalia 2nm



(f) Deuteranomalia 14nm



(g) Deuteranopia

Figura 6: Simulações de diversos níveis de deuteranomalias e protanomalias
(Fonte:[27])

2.2 SIMULAÇÃO DE VISÃO DEFICIENTE

Uma simulação da visão deficiente funciona de maneira a encontrar para cada estímulo visual cromático α , um estímulo β tal que:

- β é um dos metamerismos de α para a visão a ser simulada;
- β é o metamerismo cujas percepções do indivíduo **DVC** e do **TN** sejam o mais semelhantes possível.

Sendo uma das mais citadas referências, Brettel et al. [4] propõem um algoritmo que substitui cada cor por sua projeção em uma superfície do espaço de três dimensões de cores LMS. Esta superfície é perpendicular aos eixos de confusão dicromata. Estímulos que percorrem estes eixos são considerados metamerismos por um dicromata. A técnica apresenta resultados para dicromatas mas não funciona para os casos de tricromatas anômalos.

O algoritmo proposto por Machado et al. [27] leva em conta o processamento de otonância como modelado por Ingling e Tsouu [19]. Primeiramente são calculadas as funções de sensibilidade para cada um dos três eixos de otonância (claro-escuro, amarelo-azul e verde-vermelho) conforme as curvas de sensibilidade dos cones LMS. A Fig.8 mostra em função da frequência de luz, as curvas de sensibilidade de otonância calculadas para um sistema **TN**.

A partir destas curvas, é calculada uma matriz de transformação que leva os estímulos do espaço RGB para este espaço de otonância.

Com este modelo é possível simular protanopia, deuteranopia e tricromacia anômala de verde e vermelho abrangendo 99.34% dos casos de deficiências de visão de cores. Este modelo é o utilizado nesta dissertação como parte do cálculo do mapa de diferença de visualização entre sistemas visuais **DVC** e **TN**.



(a) Original



(b) Simulação de deuteranopia



(c) Simulação de protanopia

Figura 7: Simulações de visão deficiente utilizando o método de Brettel et al. [4]

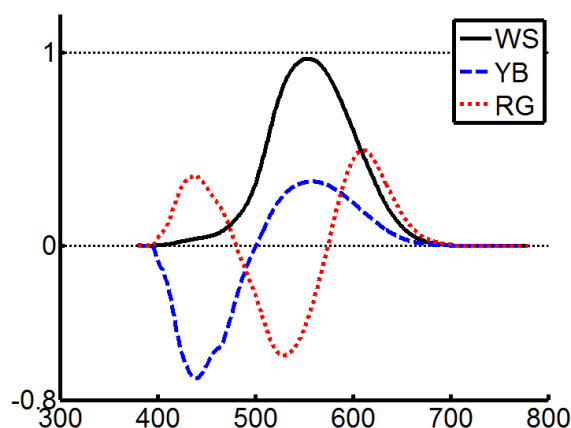


Figura 8: Sensibilidade espectral nos eixos de oponência (Fonte:[27])

2.3 RECOLORAÇÃO DE IMAGEM PARA VISUALIZAÇÃO DEFICIENTE

Diversos trabalhos tentam recolorir a imagem de maneira a facilitar a visualização da imagem por um DVC. Idealmente, uma transformação de recoloração para DVC deve permitir que a informação adquirida por um indivíduo DVC fosse idêntica à informação adquirida por TN a partir da imagem original ainda que as imagens sejam diferentes. Ou seja, o objetivo de um filtro de recoloração não é criar a mesma percepção para o indivíduo DVC, mas sim permitir que este compreenda a imagem como um indivíduo TN.

Em trabalho anterior [26] desenvolvemos um filtro que, por meio do controle de apenas uma variável, criava uma transformação linear no espaço de cores da imagem recolorindo a imagem de acordo com a necessidade do usuário. Um exemplo de sua utilização pode ser vista na Fig. 9 Esta matriz de transformação é aplicada igualmente por toda imagem e portanto mantém coerência entre diferentes quadros quando aplicado em vídeos. Implementado em GPU, tira proveito do processamento paralelo da placa. Apesar de eficiente, sua eficácia é sensível à escolha do parâmetro pelo usuário.

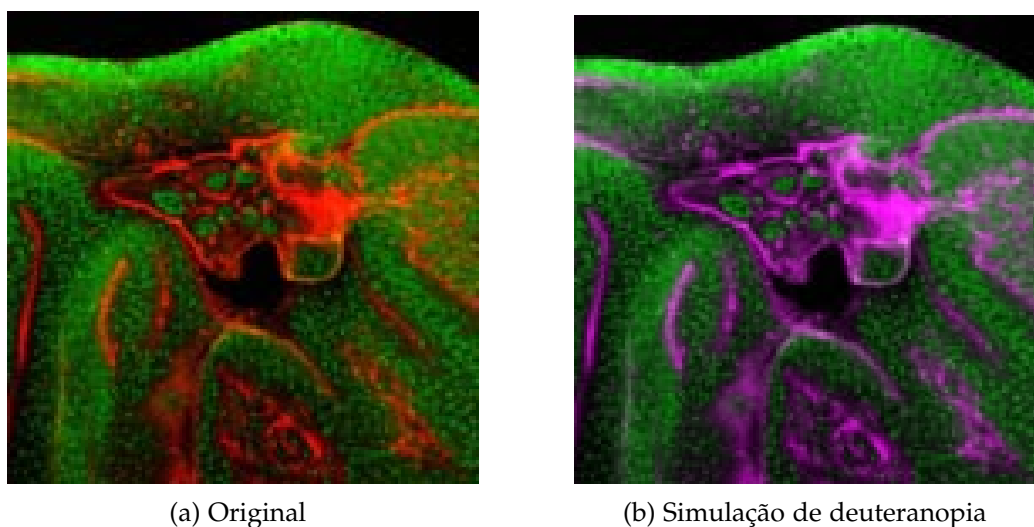


Figura 9: Imagem recolorida utilizando o método de [26]

Diversos outros trabalhos transformam a recoloração em um problema de minimização: encontrar uma função que transforme as cores da imagem minimizando uma função de erro [18, 14, 23, 14]. Estes geralmente sofrem de problemas como excesso de número de variáveis (número de cores diferentes em uma imagem) tornando-os mais lentos e com alto custo computacional. Além disso, quando aplicados em sequências de imagens, como em filmes, podem apresentar incoerências entre cada uma das imagens do conjunto uma vez que cada imagem pode ter uma solução diferente das outras.

Jefferson and Harvey [18] escolhem cores chaves como amostra para o processo minimização do erro de recoloração. O erro de coloração por sua vez é calculado a partir da diferença no espaço de cores LUV entre a cor da imagem original e a cor após a aplicação da simulação de [4]. Os resultados são interpolados para as cores não-chave.

Huang et al. [14] deixa o problema menos computacionalmente pesado diminuindo o número de amostras de cores. São criados *clusters* de cores e em seguida utiliza a distância entre os *clusters* para o processo de minimização. As distâncias entre os *clusters* são calculadas conforme a transformação de cores de TN para DVC descrita por Brettel et al. [4].

Kuhn et al. [23] tenta manter a naturalidade da imagem adicionando ao erro a distância perceptual da cor original para a obtida pela recoloração de maneira que cores que são semelhantes a tricromatas e dicromatas continuem as mesmas.

Um aspecto a ser considerado no uso destes filtros é o fato de eles alterarem a percepção da imagem. Este efeito pode ser indesejado quando a imagem vai ser visualizada por pessoas com visão normal ou em imagens cujas cores não contribuem diretamente para o entendimento da imagem. Além disso, alterar as cores de cenas naturais causa estranheza mesmo para um indivíduo com visão deficiente para cores. Sendo assim, propomos uma medida neste trabalho para avaliar a necessidade da utilização de um filtro para determinados casos.

ATENÇÃO VISUAL E SALIÊNCIA

É estimado que, por segundo, chegam o equivalente à ordem de 10^7 a 10^8 bits de informação no nervo ótico [17]. A atenção tem um importante papel na seleção das informações quebrando o problema de entendimento da cena em problemas menores.

A impressionante habilidade humana para interpretar cenas complexas com imensa quantidade de dados processados em tempo real, indica a importância da seleção de informações visuais via atenção à objetos de interesse[17][38].

3.1 ATENÇÃO VISUAL

Em seus estudos pioneiros, Yarbus [42] mostra que as trajetórias do olhar dependem tanto da tarefa realizada quanto da imagem visualizada. Seus experimentos são usados como evidência de que existe um forte relacionamento entre os pontos de fixação e seleção de atenção e pontos de interesse, relevantes à semântica da imagem. A Fig. 10 mostra os caminhos do olhar de uma pessoa ao visualizar a imagem da moça por 2 minutos. Este caminho não é uniformemente distribuído pela imagem mas sim voltado para pontos específicos como olhos e boca. A Fig. 11 mostra uma cena e na Fig. (b) o caminho realizado pelo olhar de uma pessoa ao observar a mesma cena para realizar diferentes tarefas registrado por Yarbus. Por exemplo, quando dada a tarefa de estimar a idade das pessoas na cena o olhar tende a passear mais pelas faces dos indivíduos do que quando é o observador é livre.

Alguns estímulos visuais, para um dado contexto, chamam a atenção inerentemente, ou seja, independentemente à tarefa. É o caso do vestido vermelho da garotinha no filme A Lista de Shindler na Fig. 12, ou algo

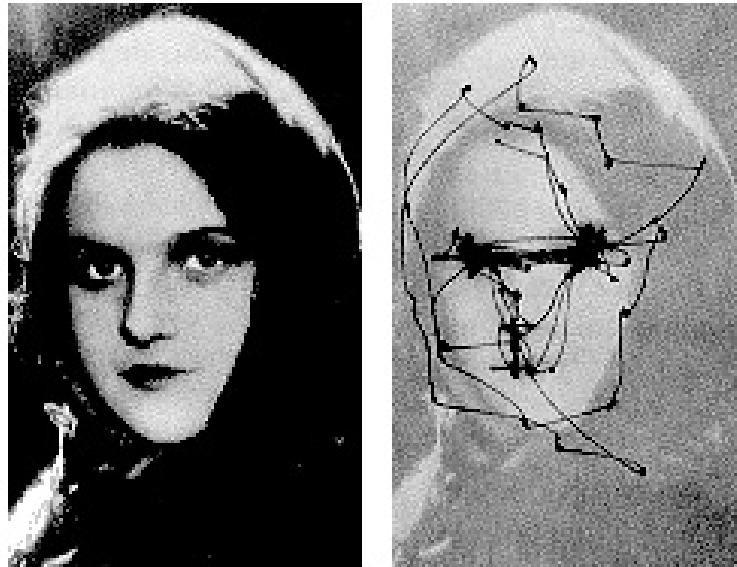


Figura 10: Movimentos do olhar de uma pessoa por 2 minutos. (Fonte:[42])

movendo-se em uma cena estática. Esta saliência é independente da natureza da tarefa realizada. Um estímulo saliente em um contexto, salta aos olhos de maneira involuntária. Esta resposta involuntária (*bottom-up*) se dá antes da atenção voluntária (*top-down*) a qual depende a tarefa a ser realizada.

A *saliência* de um ponto é a medida de ímpeto por fixação e seleção daquele ponto na imagem[40].

*Definição de
saliência de um
ponto.*

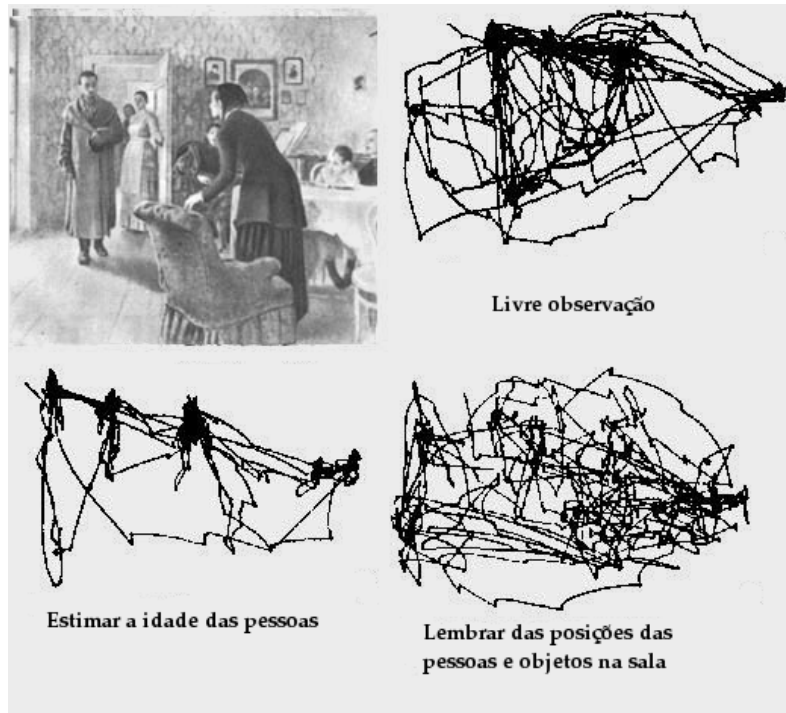
3.2 MODELOS DE SALIÊNCIA

Modelos de saliência têm sido aplicados em reconhecimento de objetos e faces[22], segmentação[34], classificação de imagens e cenas[35], *retargeting*[11]¹ e outras aplicações. Os seguintes modelos representam o estado da arte em algoritmos para prever pontos de fixação de seres humanos.

¹ *retargeting* é o processo de diminuir as dimensões de uma imagem removendo partes com menos informações



(a) Cena observada



(b) Movimentos do olhar capturados

Figura 11: (Fonte:[42])

Diversos métodos utilizam a ideia de que regiões mais dissimilares de uma imagem tornam-se mais salientes[10, 43]. Nestes trabalhos, em geral,



Figura 12: Garotinha de vermelho (Fonte: *A Lista de Shindler* (1993))

é definido algum modelo para medir as dissimilaridades entre diferentes partes da imagem e as partes mais dissimilares são consideradas as mais salientes. De maneira geral, a saliência S calculada por estes métodos poderia ser formalizada por:

$$S(x) = \sum_{\forall y \in F} D(y, x), \quad (3.1)$$

onde x e y são partes (pixels, ou janelas por exemplo) da imagem F e $D(y, x)$ é uma medida de dissimilaridade entre x e y .

Como exemplo, o modelo de saliência proposto por Zhang and Mitra [43] baseia-se na dissimilaridades entre as cores dos pixels. Neste método, os pedaços x e y seriam pixels da imagem F e a distância D seria a norma da diferença entre os vetores de cores de cada pixel multiplicada pela distância espacial entre os pixels x e y . Observando a equação 3.1 fica claro que o número de vezes que a função $D(y, x)$ precisa ser calculada é quadrático quanto ao número de partes da imagem. Sendo assim, também existe a preocupação em diminuir o número de parcelas ou de partes da imagem inclusas no cálculo. Por exemplo, o método de Zhang and Mitra [43] otimiza o cálculo agrupando pixels com cores semelhantes e utilizando a

frequência de cada cor no histograma para diminuir o número de parcelas na soma.

Existem outras possíveis abordagens para modelar a atenção visual. O modelo de Hou et al. [13] utiliza um conceito diferente, baseado em processamento de sinais, para criar um método que aproximadamente seleciona pontos de frente da imagem (*foreground*). O artigo possui uma demonstração formal de que, utilizando uma propriedade da transformada de cossenos e sua inversa, os pontos do mapa gerados pelo método com valores mais altos representam o sinal de interesse com maior probabilidade do que o fundo ou ruído.

No capítulo seguinte definimos um modelo de saliência o qual é utilizado neste trabalho para calcular as diferenças perceptuais entre diferentes sistemas de visão. Este modelo assemelha-se ao modelo proposto por [10] por usar uma transformação dos dados de maneira a codificá-los melhor e assim identificar melhor as dissimilaridades.

MODELO DE SALIÊNCIA BASEADO EM DICIONÁRIO

4.1 FRAMEWORK DE SALIÊNCIAS BASEADAS EM DICIONÁRIO DE IMAGENS

Dicionários esparsos têm sido usados com frequência para mimetizar as propriedades da estimulação do córtex visual [31] e eficientemente codificar imagens naturais[30]. A ideia básica do modelo descrito a seguir é que a imagem após ser codificada num espaço de dicionário esparsos onde suas informações são mais eficientemente descritas, a dissimilaridade de cada parte da imagem também pode ser mais eficientemente calculada.

Um dicionário esparsos, D , é um conjunto de m vetores de tamanho n cada um que formam uma base vetorial de dimensão m :

$$D = \{d_1, d_2 \dots d_m\}, d_i \in \mathbb{R}^n. \quad (4.1)$$

Ao analisarmos um conjunto de dados, um dicionário esparsos adequado para este conjunto deve permitir descrever os elementos deste conjunto de maneira eficiente, ou seja em um espaço de menor dimensionalidade. Desta maneira, a dimensão do dicionário esparsos, m é comumente menor do que a dimensão dos elementos de dados analisados, n . Na prática, a escolha de um dicionário esparsos é dependente do conjunto de dados analisados.

Dizer que um dicionário esparsos gera boas aproximações para os elementos dos dados é o mesmo que dizer que existe uma projeção no espaço gerado pelo dicionário com um erro pequeno. Ou seja, todos os vetores f de tamanho n dentre os dados analisados deve existir f' com tamanho m o qual é uma boa aproximação de f no espaço vetorial com base em D :

$$f \approx D \cdot f'. \quad (4.2)$$

Se o conjunto de dados analisados forem janelas de uma imagem maior, é possível definir um dicionário esparsos para descrever eficientemente as janelas da imagem. Usaremos a notação f_k para os valores da imagem F sob uma janela de tamanho fixo n cujo primeiro elemento é o k -ésimo pixel, F_k da imagem. Se existe um dicionário esparsos D para as janelas dessa imagem, deve seguir que:

$$f_k = D \cdot f'_k, \quad (4.3)$$

onde f'_k é chamada de projeção da janela f_k no espaço do dicionário D . Existe uma janela para cada pixel da imagem e a dimensionalidade de f'_k pode ser arbitrariamente menor do que a da janela f_k .

Definimos então o *mapa de conspicuidade visual*, S , pela distância euclidiana da codificação de cada janela para a média das codificações das outras janelas:

$$S_k = \|f'_k - \bar{f}'\|, \quad (4.4)$$

$$\bar{f}' = \frac{\sum_k f'_k}{N}, k = 1 \dots N, \quad (4.5)$$

onde N é o número de possíveis janelas dentro da imagem F e S_k é o k -ésimo pixel de S .

Resta então definirmos um método para obter D que codifique eficientemente as imagens e como fazer a projeção de cada janela no espaço gerado pela base D . A seguir definimos utilizamos a análise de componentes independentes, Análise de Componentes Independentes (ICA), para obter um dicionário esparsos e uma maneira de projetar as janelas da imagem sobre este dicionário.

4.1.1 Análise de componentes Independentes

A ICA é uma técnica estatística que procura encontrar uma representação de dimensionalidade menor para um conjunto de dados. Parte da ideia de que cada elemento dos dados originais pode ser descrito por uma combinação de diferentes componentes. Estes componentes formam uma base

de um espaço vetorial no qual os dados podem ser projetados com dimensionalidade menor do que a original. Foi verificado que os componentes do dicionário esparso gerado por essa técnica, quando aplicada a imagens naturais, assemelham-se aos campos receptores encontrados no córtex visual [31, 15, 30].

O modelo de ICA utilizado neste trabalho é o proposto por Hyvärinen e Oja [16]. Seja x a uma observação de um vetor de n variáveis aleatórias. É assumido que cada possível observação x é formada por uma combinação implícita de m fontes diferentes s . A combinação linear destas fontes pode ser descrita por uma matriz A a qual é comumente chamada de matriz de mistura:

$$x = A \cdot s = \sum_{i=1}^m a_i s_i, \quad (4.6)$$

onde a_i são as colunas de A . Podemos observar por meio desta equação e a equação 4.3 que as colunas de A são os componentes de um dicionário esparso e s é a projeção de x neste espaço.

A ICA pode ser descrita como um processo que a partir de um conjunto de observações da variável aleatória x obtém uma matriz de separação W tal que esta torna os componentes s o mais estatisticamente independentes possível. O algoritmo de ICA proposto em [16] realiza um processo de minimização da informação mútua entre os componentes resultantes da transformação de W sobre conjunto de x observados. Sendo assim, o objetivo da ICA é obter uma matriz de separação, W , a partir de diversas observações de x , tal que:

$$s = W \cdot x. \quad (4.7)$$

Para utilizarmos os componentes ICA como dicionário esparso é necessário utilizar observações de janelas para o cálculo da matriz de separação W . Em seguida, utiliza-se a matriz W para realizar a projeção de cada janela da imagem no espaço de dicionário esparso.

Nos experimentos deste trabalho, foram amostradas 1000 janelas com $n = 16 \times 16 \times 3 = 768$ elementos de cada imagem. A partir destas amostras foram calculados, usando o modelo proposto por Hyvärinen and Oja

[16], 16 componentes independentes os quais compõem um dicionário $D = \{d_1, d_2 \dots d_{16}\}$ onde $d_i \in \mathbb{R}^{768}$.

Utilizando as equações 4.4 e 4.7, o mapa de conspicuidade foi calculado da seguinte maneira:

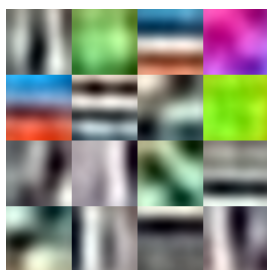
$$S_k = \left\| W \cdot f_k - \left(\frac{\sum_k W \cdot f_k}{N} \right) \right\|. \quad (4.8)$$

A Fig. 13c mostra o mapa de saliência calculado a partir da imagem original Fig. 13a com os componentes independentes da Fig. 13b.

No capítulo seguinte apresentamos uma aplicação deste modelo de saliência. A técnica foi utilizada para comparar a diferença perceptual de um sistema visual TN e um sistema DVC ao visualizar determinada imagem.



(a) Original



(b) Componentes independentes



(c) Mapa de saliência

Figura 13: Mapa de saliência calculado com componentes independentes

APLICAÇÃO DO MODELO DE SALIÊNCIA PARA CÁLCULO DA DIFERENÇA DE VISÃO NORMAL E DEFICIENTE

Neste capítulo é descrito o processo realizado para calcular a diferença perceptual para uma determinada imagem vista por sistemas de visão normal e deficiente para visão de cores. Como apontado no capítulo 3, a atenção é um importante indicador dos pontos de relevância para a compreensão da imagem. Sendo assim utilizando um modelo que simule um sistema visual deficiente é possível comparar os prováveis mapas de atenção entre os sistemas visuais.

5.1 MODELO DE VISÃO DEFICIENTE

Para modelar a visão **DVC** foram utilizados os resultados da técnica descrita por Machado et al. [27]. Esta consiste primeiramente no uso uma matriz de transformação que leva os estímulos dos cones a um espaço de oponência. Este espaço de oponência é modelado utilizando a teoria da visão de cores em dois estágios. Esta diz que após a captação pelos fotorreceptores, os estímulos devem ser comparados gerando um espaço de oponência[20] com os seguintes eixos: luz-sombra, amarelo-azul, vermelho-verde.

A técnica proposta por Machado et al. [27] modela as deficiências como alterações somente no primeiro estágio da visão - onde os fotorreceptores recebem os estímulos - sendo o segundo estágio, no espaço de oponência de cores, seria semelhante para o **TN** e **DVC**.

Com dados experimentais é calculada a matriz T_t leva as cores dos pixels do espaço RGB para este espaço de oponência do sistema visual t . Esta matriz leva em conta, para cada estímulo de cor RGB em um monitor

calibrado, a respectiva ativação de cada fotorreceptor e a posterior criação do espaço de otonência. A matriz T_t^{-1} traz de volta os elementos deste espaço de otonência ao espaço RGB. Após a aplicação de T_t , é aplicada a transformação inversa para visão TN, T^{-1} . Para simplificar a notação, omitiremos t quanto estivermos tratando um sistema TN. A matriz Γ_t corresponde à aplicação de T_t seguida de T^{-1} e é calculada pela equação 5.1. A imagem f^t que simula a imagem f vista por um sistema visual t tem seus pixels calculados pelas seguintes equações:

$$\Gamma_t = T^{-1} \cdot T_t \quad , \quad (5.1)$$

$$f_k^t = \Gamma_t \cdot f_k, \quad (5.2)$$

onde f_k é o k -ésimo pixel de f .

5.2 CALCULANDO A DIFERENÇA PERCEPTUAL ENTRE A VISÃO NORMAL E DEFICIENTE

Para realizar a comparação dos estímulos na imagem para uma visão normal e visão deficiente, propomos utilizar um modelo de saliência dada a capacidade desses modelos de selecionar os estímulos mais relevantes da imagem. Em particular utilizamos o modelo baseado em dicionário como descrito no capítulo 4.

O modelo de saliência é aplicado a partir da simulação de um determinado sistema visual para assim obter um mapa de atenção visual característico do sistema. Vamos adotar a notação de S^t como o mapa de saliência da imagem f vista por um sistema visual t o qual pode ser calculado utilizando as equações 5.2 e 4.4:

$$S_k^t = \left\| W \cdot f_k^t - \overline{W \cdot f^t} \right\|, \quad (5.3)$$

$$\overline{W \cdot f^t} = \left(\frac{\sum_k W \cdot f_k^t}{N} \right), k = 1 \dots N \quad (5.4)$$

onde $f^t k$ é a janela cujo primeiro elemento é f_k^t .

Para que seja possível comparar as magnitudes de mapas de saliências de diferentes sistemas visuais, os valores precisam ser padronizados para uma escala comum a ambos. Neste trabalho optamos pelo uso do z-valor do mapa original. O z-valor consiste no número de desvios padrão que determinado valor dista da média. Portanto, o valor de um ponto no mapa de saliência padronizado, é igual à sua distância para a média em número de desvios padrão. Desta maneira, para um determinado sistema visual t e um mapa de saliência característico deste sistema visual, S^t , dizemos que a padronização Z^t de S^t é dada por:

$$Z_k^t = \frac{S_k^t - \bar{S}^t}{\sigma(S^t)}, \quad (5.5)$$

, onde Z_k^t é o k-ésimo elemento do mapa padronizado. Com esta padronização é possível criar uma escala comum - distância à média em número de desvios padrão - a mapas de saliência oriundos de diferentes sistemas visuais.

5.2.1 Experimento

No seguinte experimento calculamos o mapa padronizado de visão com protanopia, Z^p e de visão normal, Z para diversas imagens com o fim de comparar os mapas e achar eventuais problemas de visualização nas imagens. Como no capítulo anterior, para o cálculo dos componentes independentes foram utilizadas janelas de tamanho fixo 16×16 com 3 canais de cores, ou seja, com $16 \times 16 \times 3 = 768$ elementos. Os mapas padronizados nas Fig. 14c e 14d foram subtraídos, ou seja $Z - Z^t$, e o resultado é mostrado na Fig. 14e. Desta maneira, na diferença entre os mapas de atenção, os valores mais altos significam perda de atenção naquela parte da imagem para o sistema visual t , assim como valores mais baixos significam ganho de atenção para o sistema visual t em relação à referência.

A seguir, as imagens da Fig. 15 mostram a imagem original, à esquerda e a diferença entre os mapas de atenção visual à direita para visão normal e visão protanope. Na coluna da esquerda é possível observar, em vermelho, quais as partes das imagens originais poderiam deixar de serem notadas por um indivíduo com protanopia. Com estes resultados, nosso modelo prediz uma diferença entre a atenção visual entre TN e DVC.

O modelo utilizado nos experimentos mostrou-se sensível para as variações de sistema visual deficiente. Segundo o modelo, existiria diferença de atenção entre o indivíduo com deuteranopia e protanopia com relação ao tricromata normal, como exemplificado na Fig.17.

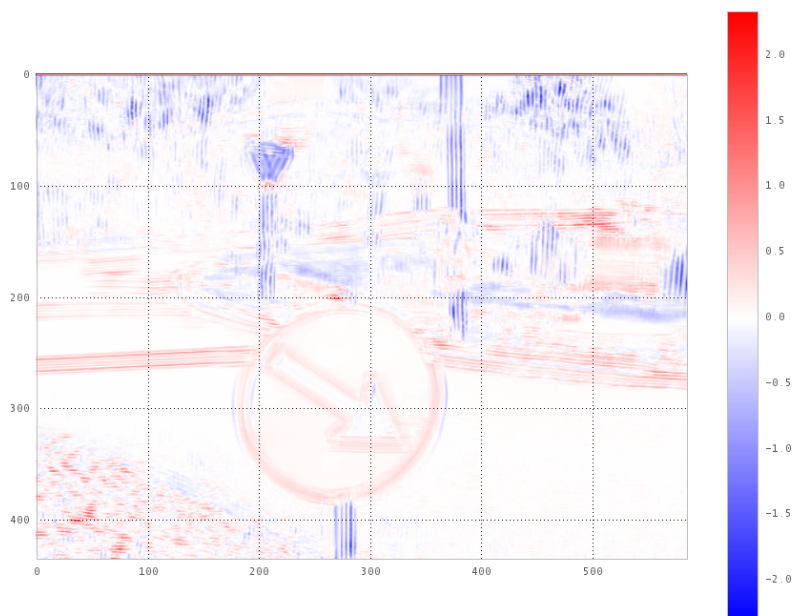
Nos experimentos cujos resultados são mostrados nas tabelas 1 e 2 comparamos os mapas de atenção das imagens originais por sua versão processada por um filtro de melhoria de visualização para DVC. Na tabela1, as imagens originais foram filtradas pela técnica descrita em [26], e as imagens da tabela 2 foram publicadas em [23] e mostram a técnica descrita no mesmo.

(a) Imagem original: f (b) Simulação de protanopia: f^t 

(c) Mapa padronizado de saliência para visão normal



(d) Mapa padronizado de saliência para visão protanope



(e) Diferença entre as saliências

Figura 14: Comparando mapas padronizados de saliência.

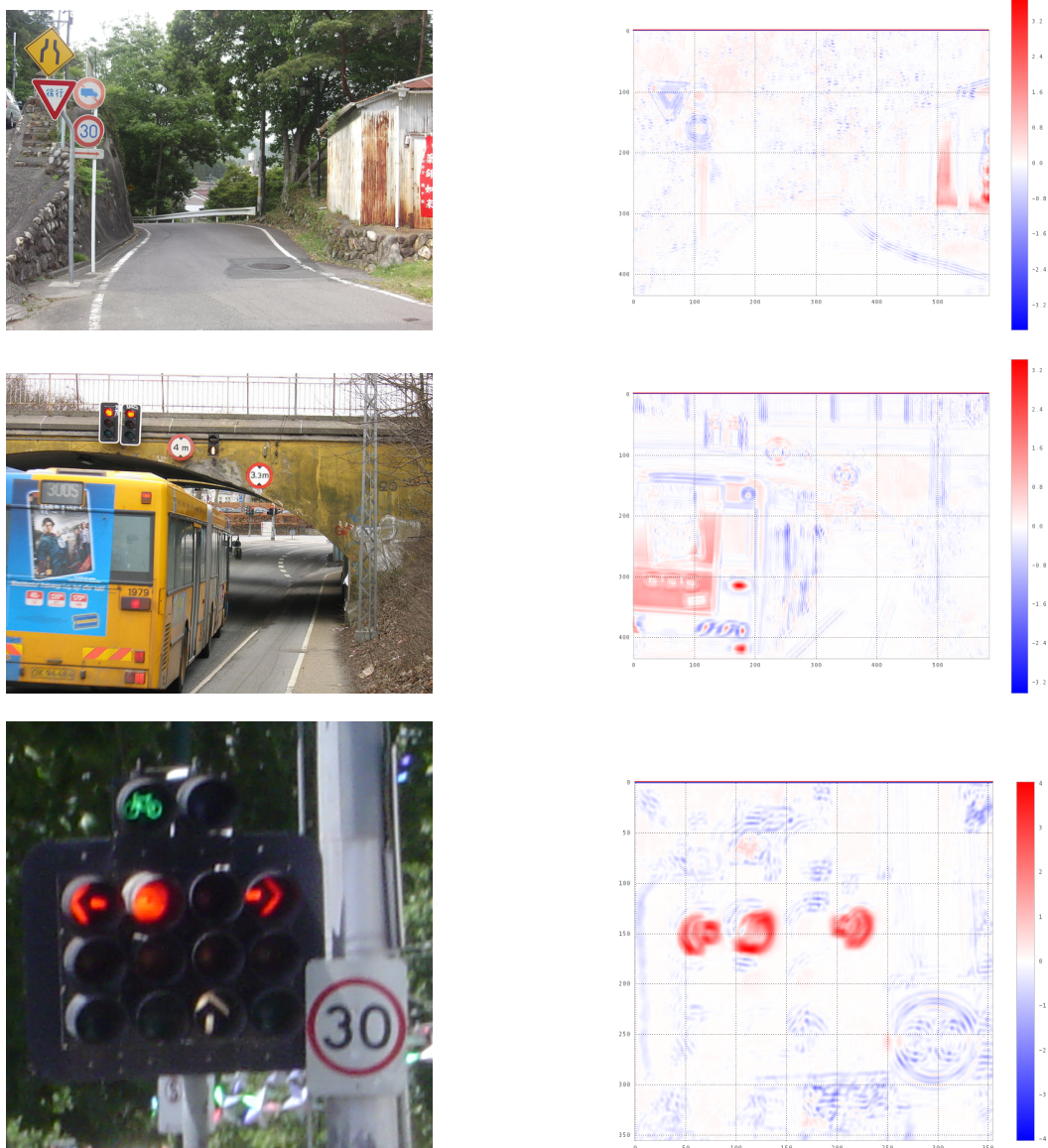


Figura 15: Diferenças perceptuais para visão normal e protranope.

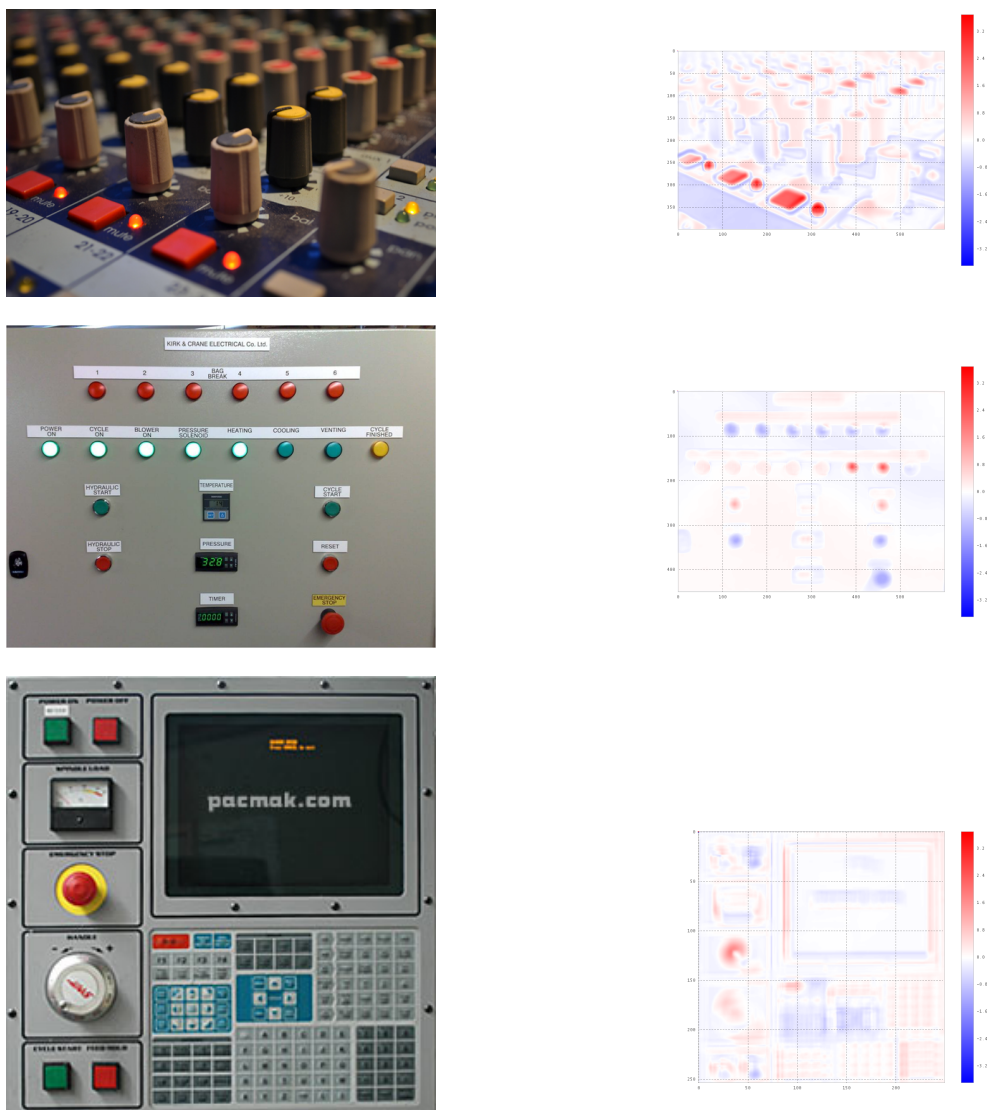


Figura 16: Diferenças perceptuais para visão normal e deuteranope.

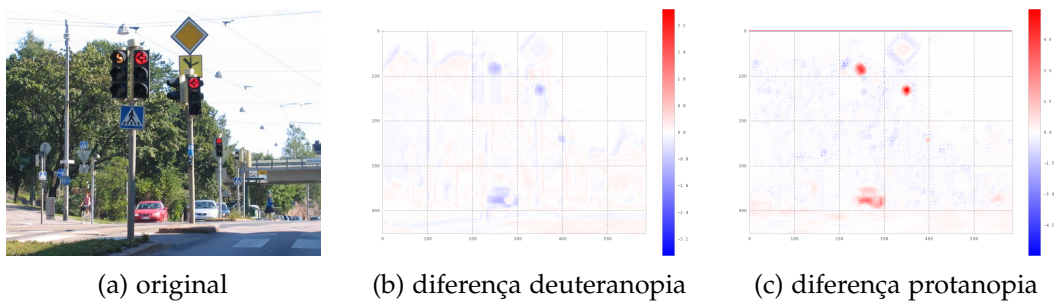


Figura 17: Diferenças de mapas normalizados para visão normal protranope e deuteranope.

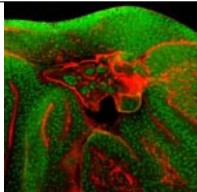
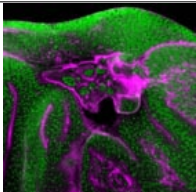
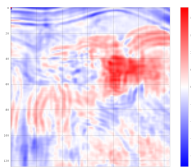
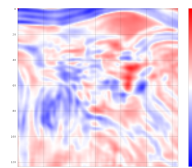
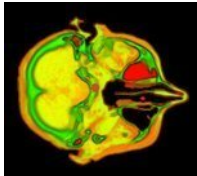
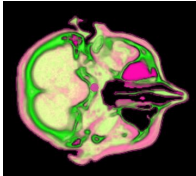
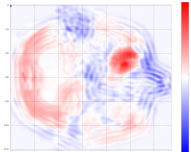
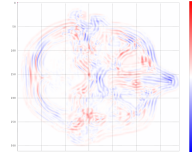
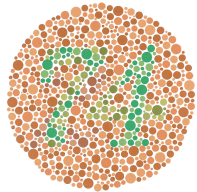
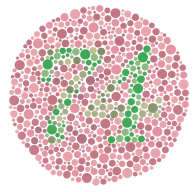
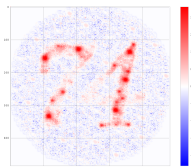
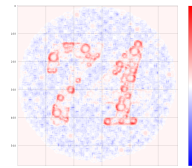
Original	Filtrada	Diferença Original	Diferença filtrada	Correlação Original	Correlação filtrada
				0.62	0.92
				0.83	0.97
				0.83	0.97

Tabela 1: Imagens filtradas pela técnica de [26] e os respectivos mapas de diferenças.


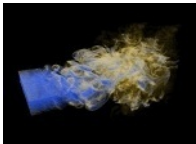
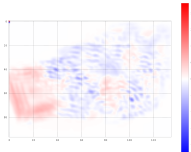
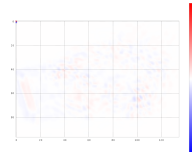
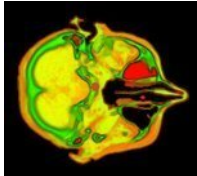
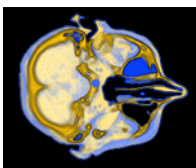
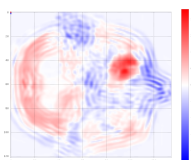
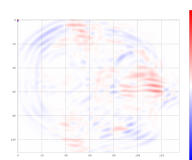


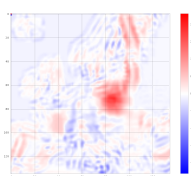
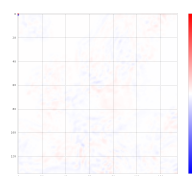


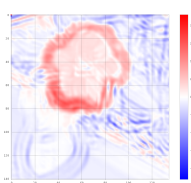
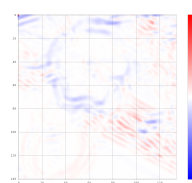
Original	Filtrada	Diferença Original	Diferença filtrada	Correlação Original	Correlação filtrada
				0.96	0.99
				0.83	0.97
				0.91	0.99
				0.83	0.99

Tabela 2: Imagens filtradas pela técnica de [23] e os respectivos mapas de diferenças.

Também foram calculados os coeficientes de correlação entre os mapas de atenção para cada uma das imagens. Em todos os casos testados a correlação entre os mapas de atenção aumenta com o uso do filtro de melhoria.

Consideramos a hipótese de que a simples diferença pixel a pixel entre a imagem original e a sua versão simulada pela visão [DVC](#) já permitiria avaliar possíveis problemas de visualização da imagem. A seguir é exibido um contra exemplo para este argumento. A [Fig. 18c](#) mostra a diferença euclidiana pixel a pixel entre a imagem original da [Fig. 18a](#) e a sua versão vista por um sistema visual protanope em [18b](#). A imagem da [Fig. 18d](#) é a diferença entre os mapas de atenção. Neste exemplo, é possível notar que a informação da diferença de atenção não necessariamente corresponde à diferença pixel a pixel.



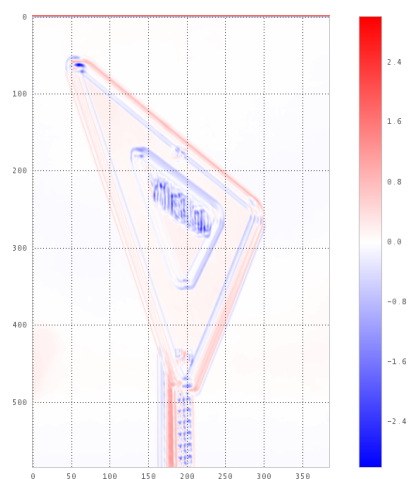
(a) Imagem original



(b) Simulação de protanopia



(c) Diferença entre pixels



(d) Diferença de atenção

Figura 18: Comparação entre a a diferença dos mapas de atenção e a diferença euclideana.

CONCLUSÃO

A métrica apresentada neste trabalho dá ensejo à mensuração da diferença perceptual das imagens entre visão anômala e normal para cores. Com isso em mãos, diferentes análises são possíveis, em especial a detecção de problemas de visualização. Isso por meio da métrica resumida por um valor único ou pela criação de um mapa de diferença; que permitiria a localização de partes da imagem com grande diferença perceptual.

A abordagem utilizada para a construção da métrica baseou-se na criação de um novo modelo computacional para saliência visual, fundamentado na análise estatística de componentes independentes, *ICA*. Este modelo pode ser generalizado num framework geral de medidas de saliência.

Dada a dificuldade em comparar mapas de saliências resultantes da análise da mesma imagem sob a ótica de diferentes sistemas visuais, neste trabalho também foi proposto um método de comparação de mapas de saliência. Este método estabelece uma padronização da escala de valores obtida em termos do número de desvio padrão em relação à media (z -valor). Por meio dessa padronização estabelece-se uma quantificação da diferença de atenção e por conseguinte da diferença de percepção. Esta quantificação consiste a métrica apresentada.

A métrica proposta foi experimentalmente testada comparando-se os resultados aferidos antes e depois da aplicação de um filtro de melhoria de imagem para deficientes. Para todas as imagens testadas, houve aumento da correlação entre os mapas de atenção, mostrando que a métrica é sensível à melhoria causada por métodos validados como o proposto por Kuhn et al. [23]. No entanto, a qualidade da aferição realizada pela métrica é dependente do modelo de saliência utilizado. Em

Por termos utilizado mapas de saliência na nossa abordagem, nos experimentos deste trabalho verificamos que, segundo nosso modelo, existe

diferença de atenção visual para a mesma cena vista por indivíduos com visão normal e por portadores de diversos tipos de deficiência. Este fato, se comprovado em testes com seres humanos, poderia ser utilizado para testar a capacidade visão de cores com pessoas com dificuldades de comunicação como bebês.

Embora, em princípio, o problema central deste trabalho raramente tenha sido explicitamente tratado em outros trabalhos (vide 1), ele é de importância fundamental para quando se deseja identificar imagens com problemas de visualização para o deficiente ou validar possíveis métodos de melhoria de visualização das mesmas. Em particular, a abordagem utilizando saliência é, potencialmente inédita para quantificar diferenças de atenção entre diferentes tipos de visão de cores.

Acreditamos que este trabalho possa ter um impacto positivo na produção de conteúdo visual cuja informação baseia-se em cores. Por tratarmos centralmente o problema de identificar e quantificar problemas de visualização para dicromatas evidenciamos que nem sempre é necessário, ou ainda desejável a aplicação de um filtro de melhoria que altere a percepção da imagem vista por um deficiente.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] D.A. Atchison, C.A. Pedersen, S.J. Dain, and J.M. Wood. Traffic signal color recognition is a problem for both protan and deutan color-vision deficient. *Hum Factors*, 45(3):495–503, 2004. (pg 1.)
- [2] J. Birch and C.M. Chisholm. Occupational colour vision requirements for police officers. *Ophthalmic Physiol Opt*, 28(6):524–31, 2008. (pg 1.)
- [3] J K Bowmaker. Visual pigments and colour vision in man and monkeys. *Journal of the Royal Society of Medicine*, 74(5):348–356, 1981. (pg vi e 9.)
- [4] Hans Brettel, Françoise Viénot, and John D. Mollon. Computerized simulation of color appearance for dichromats. *J. Opt. Soc. Am. A*, 14(10):2647–2655, Oct 1997. doi: 10.1364/JOSAA.14.002647. URL <http://josaa.osa.org/abstract.cfm?URI=josaa-14-10-2647>. (pg vi, 1, 14, 15 e 17.)
- [5] Vicente Luque Centeno, Carlos Delgado Kloos, Martin Gaedke, Martin Nussbaumer, Carlos Delgado, Wai Wcag, and Xpath Xpointer Xquery. Wcag formalization with w3c standards, 2005. (pg 1.)
- [6] B.L. Cole. The handicap of abnormal colour vision. *Clin Exp Optom*, 87(4-5):258–75, 2004. (pg 1.)
- [7] B.L. Cole. Assessment of inherited colour vision defects in clinical practice. *Clin Exp Optom*, 90(3):157–75, 2007. (pg 1.)
- [8] cvpr 2010. *The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13-18 June 2010*, 2010. IEEE. (pg 45 e 48.)

- [9] Samir S Deeb and Arno G Motulsky. Red-green color vision defects. *NCBI*, 1:1–14, 2005. URL <http://www.ncbi.nlm.nih.gov/books/NBK1301/>. (pg vi e 12.)
- [10] Antoacut;n Garcia-Diaz, Viacut;ctor Leboraacut;n, Xoseacut; R Fdez-Vidal, and Xoseacut; Pardo. On the relationship between optical variability, visual saliency, and eye fixations: A computational approach. *J Vis*, 12(6), 2012. ISSN 1534-7362. URL <http://www.biomedsearch.com/nih/relationship-between-optical-variability-visual/22693335.html>. (pg 21 e 23.)
- [11] Stas Goferman, Lihi Zelnik-Manor, and Ayellet Tal. Context-Aware Saliency Detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010. URL <http://www.ee.technion.ac.il/~ayellet/Ps/10-Saliency.pdf>. (pg 20.)
- [12] J. C. He and S. K. Shevell. Variation in color matching and discrimination among deuteranomalous trichromats theoretical implications of small differences in photopigments. *Vision Res*, 35(18):2579–88, 1995. ISSN 0042-6989. URL <http://www.biomedsearch.com/nih/Variation-in-color-matching-discrimination/7483302.html>. (pg 11.)
- [13] Xiaodi Hou, Jonathan Harel, and Christof Koch. Image signature: Highlighting sparse salient regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34:194–201, 2012. ISSN 0162-8828. doi: <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2011.146>. (pg 23.)
- [14] Jia-Bin Huang, Chu-Song Chen, Tzu-Cheng Jen, and Sheng-Jyh Wang. Image recolorization for the colorblind. In *ICASSP*, pages 1161–1164, 2009. (pg 1 e 17.)
- [15] A. Hyvarinen, P. Hoyer, and M. Inki. Topographic ica as a model of v1 receptive fields. In *Neural Networks, 2000. IJCNN 2000, Proceedings of*

- the IEEE-INNS-ENNS International Joint Conference on*, volume 4, pages 83–88 vol.4, 2000. doi: 10.1109/IJCNN.2000.860754. (pg 26.)
- [16] Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural Networks*, 13:411–430, 2000. (pg 26 e 27.)
- [17] L. Itti and C. Koch. Computational modelling of visual attention. *Nature reviews. Neuroscience*, 2(3):194–203, March 2001. ISSN 1471-003X. URL <http://view.ncbi.nlm.nih.gov/pubmed/11256080>. (pg 19.)
- [18] Luke Jefferson and Richard Harvey. Accommodating color blind computer users. In *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility, Assets '06*, pages 40–47, New York, NY, USA, 2006. ACM. ISBN 1-59593-290-9. URL <http://doi.acm.org/10.1145/1168987.1168996>. (pg vi, 1, 3, 4, 5, 6 e 17.)
- [19] Carl R. Ingling Jr. and Brian Huong-Peng Tsou. Orthogonal combination of the three visual channels. *Vision Research*, 17(9): 1075 – 1082, 1977. ISSN 0042-6989. doi: 10.1016/0042-6989(77)90013-X. URL <http://www.sciencedirect.com/science/article/pii/004269897790013X>. (pg 14.)
- [20] D B Judd. Fundamental studies of color vision from 1860 to 1960. *Proc Natl Acad Sci U S A*, 55(6):1313–30, 1966. ISSN 0027-8424. URL <http://www.biomedsearch.com/nih/Fundamental-studies-color-vision-from/5297724.html>. (pg 29.)
- [21] Deane Brewster Judd and Gunter Wyszecki. *Color in business, science, and industry / Deane B. Judd and Gunter Wyszecki*. Wiley, New York :, 3d ed. edition, 1975. ISBN 0471452122. (pg 10.)
- [22] Christopher Kanan and Garrison W. Cottrell. Robust classification of objects, faces, and flowers using natural image statistics. In *cvpr 2010 [8]*, pages 2472–2479. (pg 20.)

- [23] Giovane R. Kuhn, Manuel M. Oliveira, and Leandro A. F. Fernandes. An efficient naturalness-preserving image-recoloring method for dichromats. *IEEE Transactions on Visualization and Computer Graphics*, 14: 1747–1754, 2008. ISSN 1077-2626. doi: <http://doi.ieeecomputersociety.org/10.1109/TVCG.2008.112>. (pg [vii](#), [1](#), [6](#), [17](#), [18](#), [32](#), [38](#) e [41](#).)
- [24] Bo Liu, Meng Wang, Yichen Yang, Xiuqing Wu, and Xian-Sheng Hua. Efficient image and video re-coloring for colorblindness. In *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*, pages 906–909, 2009. doi: [10.1109/ICME.2009.5202642](https://doi.org/10.1109/ICME.2009.5202642). (pg [6](#).)
- [25] Jennifer A Long and Barbara M Junghans. Orienteers with poor colour vision require more than cunning running. *Clinical and Experimental Optometry*, 91(6):515–523, 2008. ISSN 1444-0938. doi: [10.1111/j.1444-0938.2008.00294.x](https://doi.org/10.1111/j.1444-0938.2008.00294.x). URL <http://dx.doi.org/10.1111/j.1444-0938.2008.00294.x>. (pg [1](#).)
- [26] Rafael Lopes, Andre Shoji, and Roberto Hirata. Visocor - sistema de acessibilidade para cores, 2010. URL <http://www.vision.ime.usp.br/~rafaellg/projects/visocor.html>. (pg [vi](#), [vii](#), [16](#), [17](#), [32](#) e [37](#).)
- [27] Gustavo M. Machado, Manuel M. Oliveira, and Leandro A. F. Fernandes. A physiologically-based model for simulation of color vision deficiency. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1291–1298, November/December 2009. (pg [vi](#), [1](#), [11](#), [13](#), [14](#), [16](#) e [29](#).)
- [28] J. Nathans, D. Thomas, and D.S. Hogness. Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science*, 232(4747):193–202, 1986. (pg [11](#).)
- [29] Jay Neitz and Maureen Neitz. The genetics of normal and defective color vision. *Vision Research*, 51(7):633 – 651, 2011. ISSN 0042-6989. doi: [10.1016/j.visres.2010.12.002](https://doi.org/10.1016/j.visres.2010.12.002). URL <http://www.sciencedirect.com/science/article/pii/S0042698910005699>. (pg [8](#).)

- [30] B A Olshausen and D J Field. Natural image statistics and efficient coding, 1996. (pg 24 e 26.)
- [31] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609, June 1996. ISSN 0028-0836. doi: 10.1038/381607a0. URL <http://dx.doi.org/10.1038/381607a0>. (pg 24 e 26.)
- [32] Stephen E. Palmer. *Vision science : photons to phenomenology*. MIT Press, Cambridge, Mass., 1999. (pg 10.)
- [33] A Reitner, L T Sharpe, and E Zrenner. Is colour vision possible with only rods and blue-sensitive cones? *Nature*, 352(6338):798–800, 1991. ISSN 0028-0836. URL <http://www.biomedsearch.com/nih/colour-vision-possible-with-only/1881435.html>. (pg 9.)
- [34] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. "GrabCut": interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.*, 23(3):309–314, August 2004. ISSN 0730-0301. doi: 10.1145/1015706.1015720. (pg 20.)
- [35] Ueli Rutishauser, Dirk Walther, Christof Koch, and Pietro Perona. Is bottom-up attention useful for object recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 37–44, 2004. (pg 20.)
- [36] L. T Sharpe and Karl R Gegenfurtner. *Color vision : from genes to perception / edited by Karl R. Gegenfurtner, Lindsay T. Sharpe*. Cambridge : Cambridge University Press, 1999. ISBN 0521590531. (pg 1.)
- [37] Samuel G. Solomon and Peter Lennie. The machinery of colour vision. *Nature reviews. Neuroscience*, 8(4):276–286, April 2007. ISSN 1471-003X. URL <http://dx.doi.org/10.1038/nrn2094>. (pg 8.)
- [38] John K. Tsotsos, Sean M. Culhane, Winky Yan Kei Wai, Yuzhong Lai, Neal Davis, and Fernando Nuflo. Modeling visual attention via selective tuning. *Artif. Intell.*, 78:507–545, October 1995. ISSN 0004-

3702. doi: 10.1016/0004-3702(95)00025-9. URL <http://portal.acm.org/citation.cfm?id=220069.220093>. (pg 19.)
- [39] G. WALD. The receptors of human color vision. *Science*, 145, 1964. (pg 8.)
- [40] Wei Wang, Yizhou Wang, Qingming Huang, and Wen Gao. Measuring visual saliency by site entropy rate. In *cvpr 2010* [8], pages 2368–2375. (pg 20.)
- [41] Günther Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae (Wiley Series in Pure and Applied Optics)*. Wiley-Interscience, 2 edition, August 2000. ISBN 0471399183. (pg 9.)
- [42] A. L. Yarbus. *Eye movements and vision.*, volume chapter VII. Plenum Press, New York, 1967. (pg vi, 19, 20 e 21.)
- [43] Ming-ming Cheng Guo-xin Zhang and Niloy J Mitra. Global contrast based salient region detection. In *The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition*, Colorado Springs, CO, June 2011. (pg 21 e 22.)