



FREDY ELMER ARIAS CHIPANA

UMA PROPOSTA PARA UM MODELO DE SISTEMA DE EXIBIÇÃO DE IMAGENS EM
DISPLAYS DE DISPOSITIVOS MÓVEIS BASEADO EM UM MÉTODO DE ATENÇÃO VISUAL

CAMPINAS

2014



UNIVERSIDADE ESTADUAL DE CAMPINAS

Faculdade de Engenharia Elétrica e de Computação

Fredy Elmer Arias Chipana

UMA PROPOSTA PARA UM MODELO DE SISTEMA DE EXIBIÇÃO DE IMAGENS EM
DISPLAYS DE DISPOSITIVOS MÓVEIS BASEADO EM UM MÉTODO DE ATENÇÃO VISUAL

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica.
Área de concentração: Telecomunicações e Telemática.

Orientador: Prof. Dr. Yuzo Iano

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL
DA DISSERTAÇÃO DE MESTRADO DEFENDIDA PELO
ALUNO FREDY ELMER ARIAS CHIPANA, E
ORIENTADA PELO PROF. DR. YUZO IANO.

CAMPINAS

2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Luciana Pietrosanto Milla - CRB 8/8129

Ar41p Arias Chipana, Fredy Elmer, 1979-
Uma proposta para um modelo de exibição de imagens em displays de dispositivos móveis baseado um método de atenção visual / Fredy Elmer Arias Chipana. – Campinas, SP : [s.n.], 2014.

Orientador: Yuzo Iano.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Análise de imagem. 2. Visualização. 3. Imagens. I. Iano, Yuzo, 1950-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: A proposal for a model of system for exhibition of images in displays of mobile devices based on a method of visual attention

Palavras-chave em inglês:

Image analysis

Visualization

Images

Área de concentração: Telecomunicações e Telemática

Titulação: Mestre em Engenharia Elétrica

Banca examinadora:

Yuzo Iano [Orientador]

Evaldo Gonçalves Pelaes

Vicente Idalberto Becerra Sablón

Data de defesa: 25-07-2014

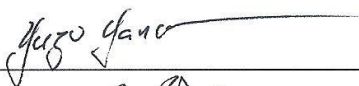
Programa de Pós-Graduação: Engenharia Elétrica

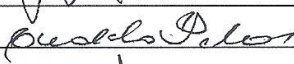
COMISSÃO JULGADORA - TESE DE MESTRADO


Candidato: Fredy Elmer Arias Chipana

Data da Defesa: 25 de julho de 2014

Título da Tese: "Uma Proposta para Um Modelo de Sistema de Exibição de Imagens em Displays de Dispositivos Móveis Baseado em Um Método de Atenção Visual"

Prof. Dr. Yuzo Iano (Presidente): 

Prof. Dr. Evaldo Gonçalves Pelaes: 

Prof. Dr. Vicente Idalberto Becerra Sablón: 

RESUMO

A apresentação de imagens em telas de dispositivos móveis tem limitações que dependem da experiência do usuário. A adaptação de imagens consiste em acondicionar o tamanho e a resolução para obter uma visualização na tela do dispositivo. O uso de mecanismos de atenção visual permite reduzir esforços no processamento do estímulo visual do olho humano. Os modelos de atenção visual também ajudam a reduzir a complexidade computacional em aplicações de processamento de imagens. Propõe-se neste trabalho uma melhoria de um modelo de atenção visual que é aplicado à adaptação de imagens em telas de dispositivos móveis.

Palavras chave: adaptação de imagens, modelos de atenção visual, mapas de relevância.

ABSTRACT

The presentation of images on screens of mobile devices has limitations that depend on the user experience. The adaptation of images is pack the size and resolution in order to visualize the images on the device screen. The use of a visual attention mechanism allows reducing efforts in processing the visual stimulus of the human eye. Models of visual attention also help to reduce the computational complexity in image processing applications. We propose an improvement to this model of visual attention to be applied to the adaptation of images on screens of mobile devices.

Keywords: images adaptation, visual attention models, maps of saliency.

SUMÁRIO

RESUMO.....	vii
<i>ABSTRACT</i>	vii
SUMÁRIO.....	ix
AGRADECIMENTOS.....	xi
LISTA DE FIGURAS.....	xv
LISTA DE SÍMBOLOS.....	xxi
CAPÍTULO 1 - INTRODUÇÃO.....	1
1.1 Motivação.....	3
1.2 Caracterização do problema.....	4
1.2.1 Objetivo principal.....	6
1.2.2 Objetivos secundários.....	6
1.3 Definições.....	6
1.3.1 Mecanismos de Seleção.....	7
1.3.2 Atenção visual e relevância.....	7
1.4 Modelos existentes.....	8
1.4.1 Modelos de adaptação de imagem.....	8
1.4.2 Modelos de atenção visual.....	8
1.5 Relevância dos modelos estudados.....	20
1.6 Organização da tese.....	20
CAPÍTULO 2 - DESCRIÇÃO DO MODELO PROPOSTO.....	21
2.1 Descrição do modelo de atenção visual.....	21
2.2 Atenção visual seletiva.....	23
2.2.1 Atenção visual utilizando relevância.....	24
2.2.2 Atenção visual utilizando a face.....	26
2.2.3 Atenção visual utilizando um texto.....	28
2.2.4 Combinação dos mapas de atenção.....	29
2.3 Adaptação de imagens.....	30
2.3.1 Fidelidade de informação.....	30
2.3.2 Algoritmo de adaptação.....	31
2.3.3 Conjunto de objetos de atenção válidos.....	33

2.3.4	Conjunto ótimo de objetos de atenção.....	34
2.3.5	Adaptação da solução na região objetivo	34
CAPÍTULO 3 - METODOLOGIA		36
3.1.1	Distância de Mahalanobis.....	36
CAPÍTULO 4 - RESULTADOS.....		38
CAPÍTULO 5 - CONCLUSÕES E TRABALHO FUTURO		41
5.1	Análise dos resultados e comentários gerais	42
5.2	Contribuições.....	42
REFERÊNCIAS.....		43
Apêndice 1 - Fluxogramas dos programas.....		47
Apêndice 2 - Códigos com detalhamentos.....		48
Apêndice 3 - Resultados		50
Apêndice 4 - Resultados visuais		61

AGRADECIMENTOS

Gostaria de agradecer a todos aqueles que me apoiaram nestes anos de pesquisa e aprendizagem. Primeiramente, a toda minha família pelo apoio moral durante esta jornada. Ao meu orientador, Prof. Yuzo Iano, sou grato pela orientação durante esses meses. Ao meu colega e amigo Roger Larico por seus conselhos e conhecimentos que serviram para definir meu projeto. Aos colegas do Laboratório de Comunicações Visuais, pela amizade, companheirismo e sugestões para este trabalho. Aos professores e funcionários do Programa de Pós Graduação em Engenharia Elétrica da Universidade Estadual de Campinas, por sua ajuda sempre necessária. A CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), a FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), a FAEPEX-UNICAMP (Fundo de Apoio ao Ensino, à Pesquisa e Extensão) e ao SAE (Serviço de Apoio ao Estudante) pelo apoio financeiro.

A todos que, direta ou indiretamente, colaboraram para a realização deste trabalho.

“Se uma pessoa é perseverante, por mais que tenha falta de entendimento, se tornará inteligente e mesmo que fracasse se transformará em alguém forte.”

-Leonardo da Vinci

LISTA DE FIGURAS

Figura 1. Exemplo de adaptação.	2
Figura 2. Exemplo de atenção visual de <i>software online</i>	4
Figura 3. Caracterização do problema de adaptação (imagem tomada do banco de dados) [44].	5
Figura 4. Diagrama esquemático usado para modelos cognitivos [25].	9
Figura 5. Exemplo de modelo bayesiano usado em [31].	10
Figura 6. Filtro de retina adaptável.	12
Figura 7. Modelo de informação teórica baseado na entropia.	13
Figura 8. Exemplo de modelo gráfico usado em [35].	15
Figura 9. Exemplo de modelo usando a transformada de Fourier usado em [31].	16
Figura 10. Exemplo de padrão usado em [36].	17
Figura 11. Exemplo de classificação de padrões [36].	18
Figura 12. Exemplo de mapas de relevância dos modelos estudados em [15].	19
Figura 13. Diagrama de fluxo do modelo proposto.	21
Figura 14. Objetos de atenção em uma cena. A imagem foi tomada de um banco de dados [44].	22
Figura 15. Função gaussiana para a relevância da cor.	25
Figura 16. Exemplo da atenção visual usando um algoritmo de relevância.	25
Figura 17. Esquema geral de reconhecimento de faces.	26
Figura 18. Exemplo de características básicas usadas em [19].	27
Figura 19. Exemplo de detecção de face.	28
Figura 20. Exemplo de detecção de texto.	29
Figura 21. Possível conjunto de 3 objetos de atenção.	35
Figura 22. Gráfico de dispersão dos resultados para o modelo base original e para o banco de dados. Valores médios.	38
Figura 23 Gráfico de dispersão dos resultados para o sistema proposto e para o banco de dados.	39

LISTA DE TABELAS

Tabela 1. Comparação das distâncias calculadas.....	39
---	----

LISTA DE ABREVIATURAS E SIGLAS

AO	<i>Attention Object</i>
AV	<i>Attention Value</i>
BD	Banco de Dados
BU	<i>Bottom-Up</i>
CRF	<i>Conditional Random Fields</i>
DBN	<i>Dynamic Bayesian Networks</i>
GBVS	<i>Graph-Based Visual Saliency</i>
GPU	<i>Graphics Process Unit</i>
HMM	<i>Hidden Markov Models</i>
MP	Modelo Proposto
MPS	<i>Minimum Perceptible Size</i>
RGB	<i>Red, Green, Blue</i>
ROI	<i>Region Of Interest</i>
SVM	<i>Support Vector Machine</i>
TD	<i>Top-Down</i>

LISTA DE SÍMBOLOS

L_C	Mapa de cor
L_I	Mapa de intensidade
L_o	Mapa de orientação
c	Escala do centróide
N	Operador de normalização
s	Escala da envoltória
ι	Característica
S	Matriz de covariância
D_M	Distância de Mahalanobis

CAPÍTULO 1 - INTRODUÇÃO

O crescente avanço das tecnologias nas áreas de computação e comunicações permite que os dispositivos eletrônicos possam acessar e interagir usando informações de diferentes tipos e mídias.

Essa interação apresenta limitações que dependem da experiência do usuário em acessar o conteúdo de mídias acondicionado ao tamanho e à resolução da tela. Para se visualizar as imagens que possuem tamanho maior do que as telas dos dispositivos, cortes ou redimensionamentos são necessários a fim de adaptar as imagens aos formatos diferentes de telas dos dispositivos. Nesse contexto, encontram-se muitas propostas como em [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. Tais propostas são baseadas principalmente no uso de um modelo de atenção visual.

Os modelos de atenção visual visam se aproximar do modelo da percepção visual humana. A percepção visual baseia-se em uma organização mental. A interpretação da informação sensorial visual pelo cérebro permite alcançar uma consciência e compressão do ambiente local. Dessa forma, por exemplo, pode-se notar a existência de objetos e assimilar eventos gerando conhecimento. O conhecimento é a faculdade humana de processar essa informação. Esse processo envolve atenção, expectativa, aprendizagem, memória, linguagem e resolução de problemas [11].

A partir de um único estímulo visual não é possível processar totalmente a informação de imediato. Primeiro é necessário entender esse estímulo. A complexidade envolvida pode ser grande e por isso a atenção visual é uma área de pesquisa muito promissora para se processar, segmentar, e hierarquizar uma informação.

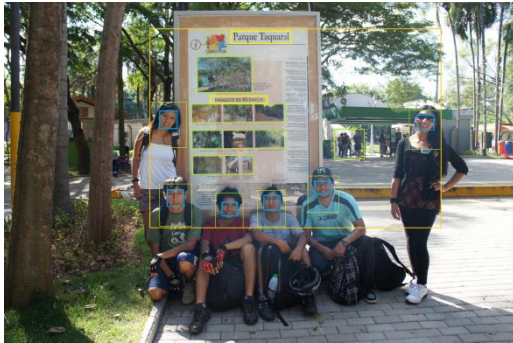
No contexto computacional, uma determinada região de uma cena pode receber uma maior atenção visual do que outras regiões. Na verdade, algumas dessas regiões podem inclusive serem ignoradas. Desse modo, este trabalho visa direcionar um enfoque na atenção visual para adaptação de uma imagem em uma região menor ou mesmo diferente, buscando conservar, na medida do possível, o mesmo impacto visual. A Figura 1 mostra exemplos de adaptação de uma imagem onde a informação da cena é substituída com pouca perda de informação visual. Nesses casos, as faces das pessoas e o texto são considerados de maior importância.



a. Imagem original.



b. Detecção de faces e texto.



c. Atenção visual (no original).



d. Recorte da imagem.

Figura 1. Exemplo de adaptação.

Para a proposta deste trabalho, serão aplicados algoritmos de detecção de relevância [17], texto [20] e faces [19]. Devem-se determinar as regiões de maior importância, buscando-se adaptar a imagem em diferentes formatos de telas.

Para visualizar uma imagem em um dispositivo móvel (celular, *smartphones*, *tablets* e *notePCs*) que tenha uma tela de menor tamanho e resolução, é necessário enviar uma informação visual útil para o dispositivo (usuário). Desse modo, será possível obter vantagens quantitativas tais como, menor taxa de transmissão de dados e regiões de interesse baseadas no modelo de atenção visual. Além disso, será possível obter vantagens subjetivas como melhor visualização em telas onde é necessário fazer uma ampliação ou *zoom* para se enxergar detalhes.

1.1 Motivação

A construção de sistemas capazes de trabalhar em tempo real tais como visão por computador ou orientação de robôs é uma tarefa que exige poder computacional. se busca um mecanismo de seleção baseado em níveis de relevância usados pelos humanos. Os olhos recebem uma cadeia de estímulos de dados visuais (108-109 *bits*) a cada segundo aproximadamente e a atenção é facilitada pela retina que tem uma resolução maior na fóvea central e uma resolução menor na periferia em termos de quantidade de células sensoriais (cones e bastonetes).

Em décadas recentes, muitos cientistas enfocaram seus estudos em comportamentos relacionados à atenção visual e muitos outros revisaram modelos de atenção psicológicos, neurobiológicos e de perspectiva computacional. Alguns modelos de atenção visual foram propostos inspirados em estudos que levavam em conta mecanismos para reduzir complexos processos tais como, o reconhecimento de objetos ou tarefas de interpretação de cenas ([1], [2], [3], [4], [5], [6], [7], [8], [9], [10]).

Os dispositivos móveis, em geral, são projetados para apresentarem uma visualização rápida a curtas distâncias. A partir disso, o sistema de comunicações móveis permite acesso ubíquo ou paralelo a conteúdo multimídia, tornando a recepção dos sinais independente dos tamanhos dos dispositivos. Devido às limitações intrínsecas aos diferentes formatos de telas, um importante passo é adaptar a visualização de acordo com as necessidades do usuário em diferentes ambientes.

A diversidade de tipos de usuários, de conteúdo multimídia e de terminais gera a necessidade de desenvolvimento de soluções que permitam um acesso universal da informação. Isso significa que as ferramentas de adaptação devem permitir apresentações da mesma informação, se adequando às diferentes condições de uso.

Existem muitas aplicações tecnológicas para o uso desses modelos, como filtragem por propriedades de estímulos, procura visual, identificação de texto manuscritos, localização de objetos, visão de robôs, codificação de imagens, inspeção automática de imagens, etc.

Inicialmente, neste trabalho, é descrito um modelo de atenção usado baseado em [19], definindo-se os objetos de atenção obtidos do modelo de atenção visual. Para resolver esse problema de adaptação é levado em conta um arranjo desses objetos buscando-se um arranjo o mais próximo possível do objetivo. Para tanto, usa-se um método de maximização de valor de atenção descrito no próximo capítulo. Assim é criado um esquema de adaptação automática para envio seletivo de imagens para dispositivos móveis. Em seguida, são apresentadas as simulações e a discussão final sobre o desempenho do sistema.

Existem *softwares online* que oferecem serviços de atenção visual especialmente focado no *marketing* (Figura 2. Exemplo de atenção visual de *software online*. Figura 2). Apresentam-se imagens correspondentes nessa figura para visualização. As regiões consideradas importantes são marcadas em *high light* ou tons diferentes.



a. Exemplos de Feng-GUI. (<http://www.feng-gui.com/>).

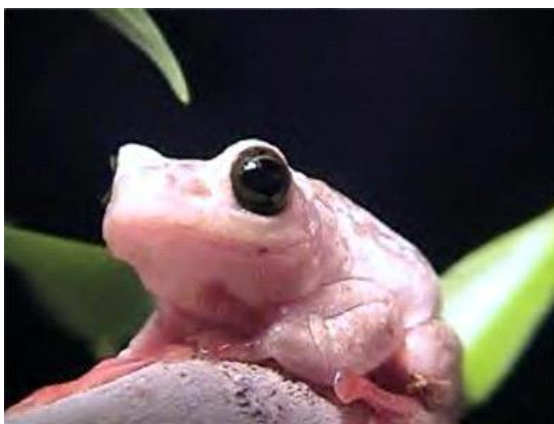


b. Exemplos de 3M. (http://solutions.3m.com/wps/portal/3M/en_US/VAS-NA/VAS/).

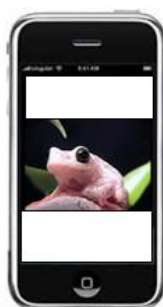
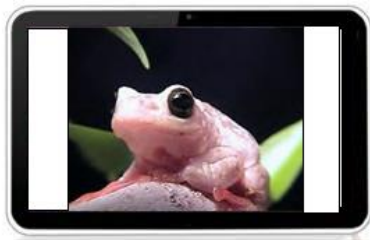
Figura 2. Exemplo de atenção visual de *software online*.

1.2 Caracterização do problema

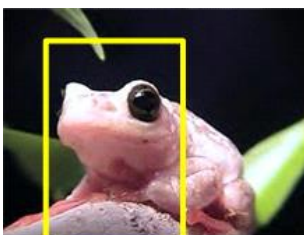
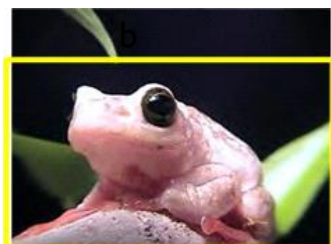
Adaptar uma imagem dinamicamente consiste em modificar seu conteúdo para que corresponda ao tamanho da tela do dispositivo alvo preservando as regiões consideradas relevantes. Geralmente, o que se faz é mudar o tamanho da imagem. Nesse caso, quando se usa um modelo de atenção visual, o que se quer é determinar uma região de interesse e em seguida realizar um corte nessa imagem que corresponda à tela desejada salvando a importância visual da imagem ainda que haja uma perda de resolução. Neste trabalho leva-se em conta que para a percepção humana são importantes a presença de rostos/faces ou textos. Assim, esses últimos são considerados importantes para preservação.



a. Imagem original.



b. Imagem tradicionalmente apresentada.



c. Adaptação da imagem com o modelo de atenção visual.



d. Adaptação da imagem no dispositivo.

Figura 3. Caracterização do problema de adaptação (imagem tomada do banco de dados) [44].

A Figura 3. Mostra visualmente o problema. Nessa figura, a imagem original (a) é adaptada de maneira tradicional em (b) deixando espaços em branco na tela do dispositivo. Em (c) mostra-se um redimensionamento e seleção da área de atenção visual ideal para a tela objetivo. Em (d) mostra-se a solução de adaptação com um modelo de atenção visual.

1.2.1 Objetivo principal

O objetivo principal deste trabalho é propor um modelo de sistema de exibição de imagens em *displays* de dispositivos móveis, onde a imagem alvo seja redimensionada e segmentada para adaptar-se a uma tela menor e de diferente formato com perda de percepção visual reduzida.

1.2.2 Objetivos secundários

Por outra parte propõe-se um modelo de atenção visual, o qual está baseado no modelo usado em [19]. Esse método escolhido proporciona um mapa de atenção visual geral para ser usada como parâmetro de medida de nível de atenção. Esse modelo é composto de três partes: uma, para a saliência básica (cor, textura e orientação), a segunda de um detector de faces e por último um detector de texto. Mais adiante se especifica a contribuição do presente trabalho.

1.3 Definições

A atenção é um conceito geral que cobre todos os fatores que influenciam os mecanismos de seleção. O olhar fixo necessita de movimentos coordenados de olhos e cabeça e é usado como uma forma de procurar algum objeto ou comportamento natural que chame a atenção [20].

Muitos fatores podem determinar quais as características da imagem que devem ser selecionadas ou descartadas no processo de atenção. Esse processo de seleção tem duas categorias [35], descritos em seguida.

1.3.1 Mecanismos de Seleção

- **Top-down (BU):** depende do estado interno do sujeito em que sua atenção é uma função de motivação para análise de uma cena. Nesse caso, são usados detecção de faces e de texto.
- **Bottom-up (TD):** onde o sujeito não procura um objeto específico. De fato, sua atenção é dirigida aos diferentes elementos da cena de maneira inconsciente, denominados também de relevância [21] da cena. A relevância de uma imagem ou cena intuitivamente é formada por dois mecanismos, sendo um deles a simples medida de visibilidade de uma localização na cena e o outro, a seleção de unidades com maior valor de visibilidade.

1.3.2 Atenção visual e relevância

Os termos atenção e relevância são muitas vezes usados como sinônimos. No entanto, cada um desses dois termos tem uma definição mais sutil que permite a sua delimitação.

A atenção é um conceito geral que abrange todos os fatores que influenciam os mecanismos de seleção, sejam eles cenas *bottom-up* (BU) ou orientada à expectativa *top-down* (TD).

A relevância caracteriza intuitivamente algumas partes de uma cena. Essas partes poderiam ser objetos ou regiões que aparecem ao observador como sobressaindo em relação às suas partes vizinhas. O termo "relevante" é muitas vezes considerado no contexto de cálculos *bottom-up* [22], [25].

A relevância em uma região é determinada principalmente pelas diferenças existentes quanto a sua cor, orientação, movimento, profundidade, etc. [11].

1.4 Modelos existentes

Nesta secção são descritos brevemente, os modelos de adaptação e de atenção visual encontrados na literatura.

1.4.1 Modelos de adaptação de imagem

Os modelos de adaptação de imagens variam muito de acordo com as tarefas que desempenham. Geralmente, o que se realiza nesses diferentes métodos é uma redução ou ampliação da imagem que corresponda ao tamanho do dispositivo de reprodução. Como por exemplo, o algoritmo de A. Saif [30] que está baseado na proximidade de *pixels* vizinhos mantendo-se a textura principal.

1.4.2 Modelos de atenção visual

Os modelos são citados a seguir. A base de seus mecanismos é da obtenção de relevância. Alguns modelos podem fazer parte de mais de uma categoria. Esses modelos são apresentados em ordem cronológica.

Modelos cognitivos

A maioria dos modelos de atenção é diretamente ou indiretamente inspirada em conceitos cognitivos.

Modelos básicos como de Itti [25] usam características como cor, intensidade e orientação. Essa é a base de outros modelos também. Uma imagem é subamostrada em uma pirâmide gaussiana. Cada nível da pirâmide é descomposto em faixas de vermelho (R), verde (G), azul (B), amarelo (Y), intensidade (I) e orientações locais (O). A partir do centro dessas faixas é construído e normalizado o mapa de características f . Cada característica é computada por um conjunto de operações lineares da envoltória de centro que se aproximam do campo visual receptivo. Tipicamente, os neurônios são mais sensitivos em pequenas regiões do espaço visual (no centro) e menos na região concêntrica antagônica (na envoltória) [43].

A Figura 4 amostra um diagrama esquemático básico usado em modelos cognitivos onde a imagem é filtrada por cor, intensidade e orientação. Em seguida criam-se mapas de relevância de cada um deles que são combinados em um único mapa. Esse mapa representa o valor de atenção visual para a imagem.

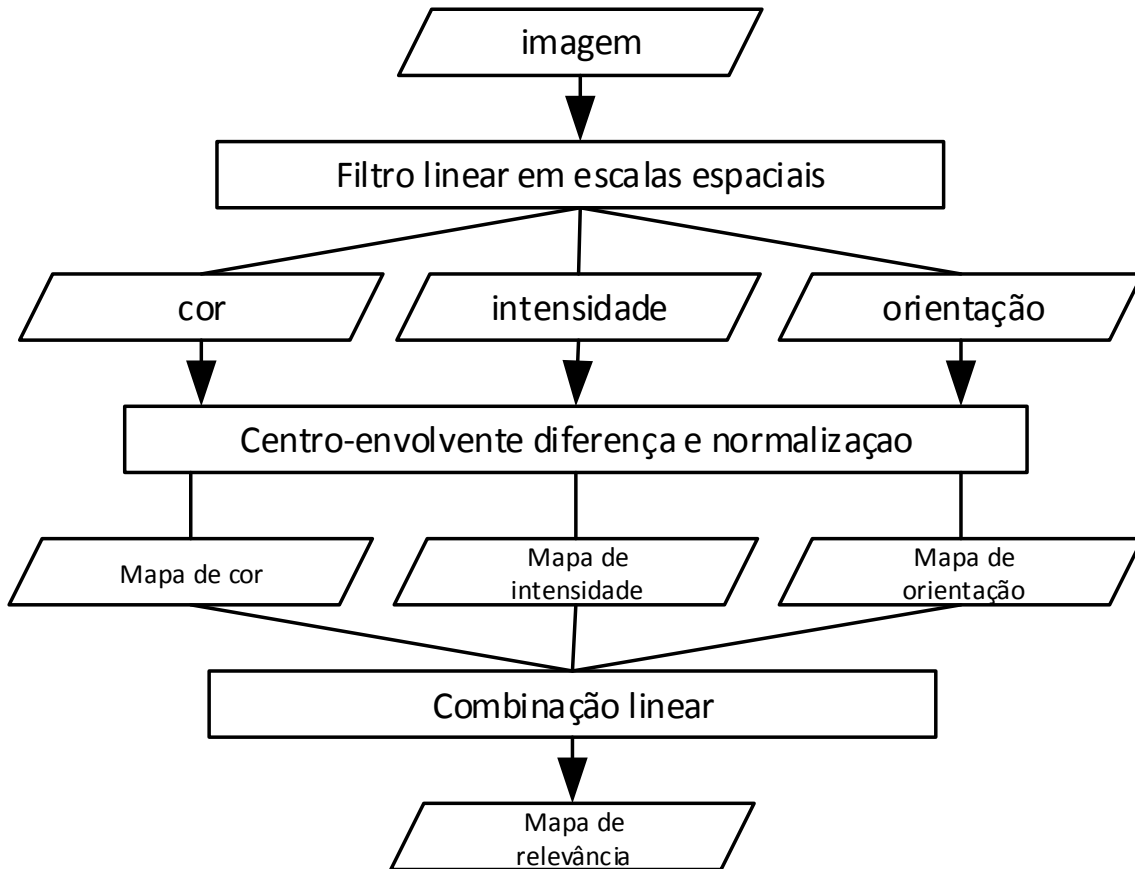


Figura 4. Diagrama esquemático usado para modelos cognitivos [25].

Modelos cognitivos têm a vantagem de expandir a nossa visão de bases biológicas da atenção visual. Isso tem ajudado a entender ainda mais os princípios computacionais, bem como contribuído para um entendimento dos mecanismos neurais deste processo e de outros processos dependentes complexos, relacionados com reconhecimento de objetos.

Modelos bayesianos

Esses modelos usam uma combinação probabilística. Eles combinam evidência sensorial com as restrições dadas. Por exemplo, o modelo de Zhang [31] envolve cenas dinâmicas através da introdução de filtros temporais (diferença de exponenciais) e usa uma montagem com distribuição de Gauss generalizada com uma distribuição estimada para cada resposta do filtro. O modelo calcula suas características e estima a relevância *bottom-up* para cada ponto. Usam filtros concebidos para serem eficazes e semelhantes ao do sistema visual humano. As distribuições de probabilidade dessas características espaço-temporais foram apreendidas a partir de um conjunto de vídeos de ambientes naturais.

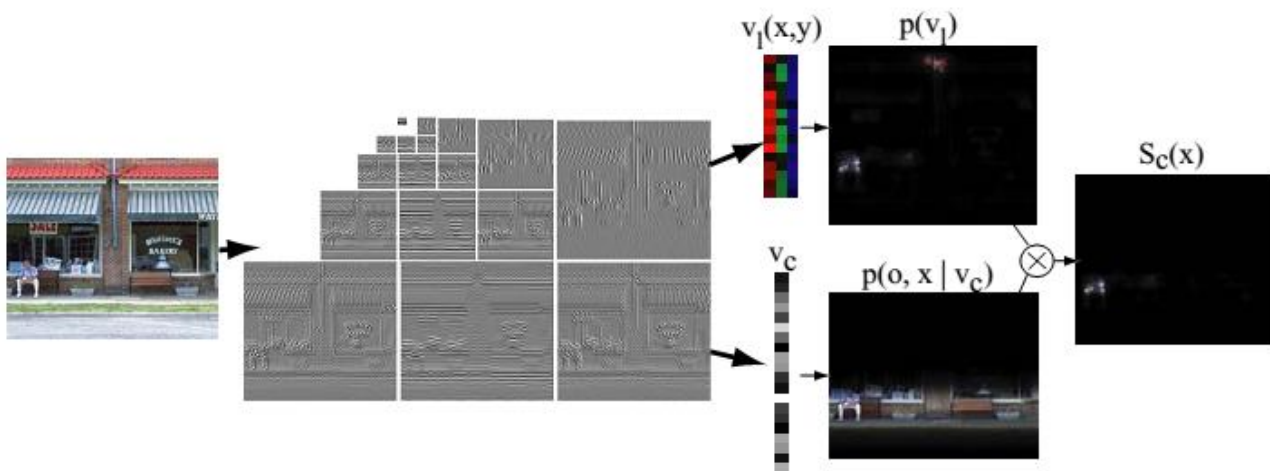


Figura 5. Exemplo de modelo bayesiano usado em [31].

A Figura 5 mostra visualmente o esquema usado por [31] para um sistema de detecção de objetos integrando relevância e prioridade. As equações 1 e 2 são usadas por esse modelo. A figura mostra a imagem original e a filtragem de características. Aplicando-se o teorema de Bayes, o autor obtém $S_c(x)$ que determina a relevância local para esse modelo.

$$P(o, x|V_I, V_C) = \frac{P(V_I|x, V_C)}{P(V_I|V_C)} P(o, x|V_C) \quad (1)$$

$$S_c(x) = \frac{P(o, x|V_C)}{P(V_I|V_C)} \quad (2)$$

Onde:

$V_I(x, y)$ é o vetor de características inicial;

$p(V_I)$ é a probabilidade de encontrar essa característica inicial;

V_C é o vetor contextual de características.

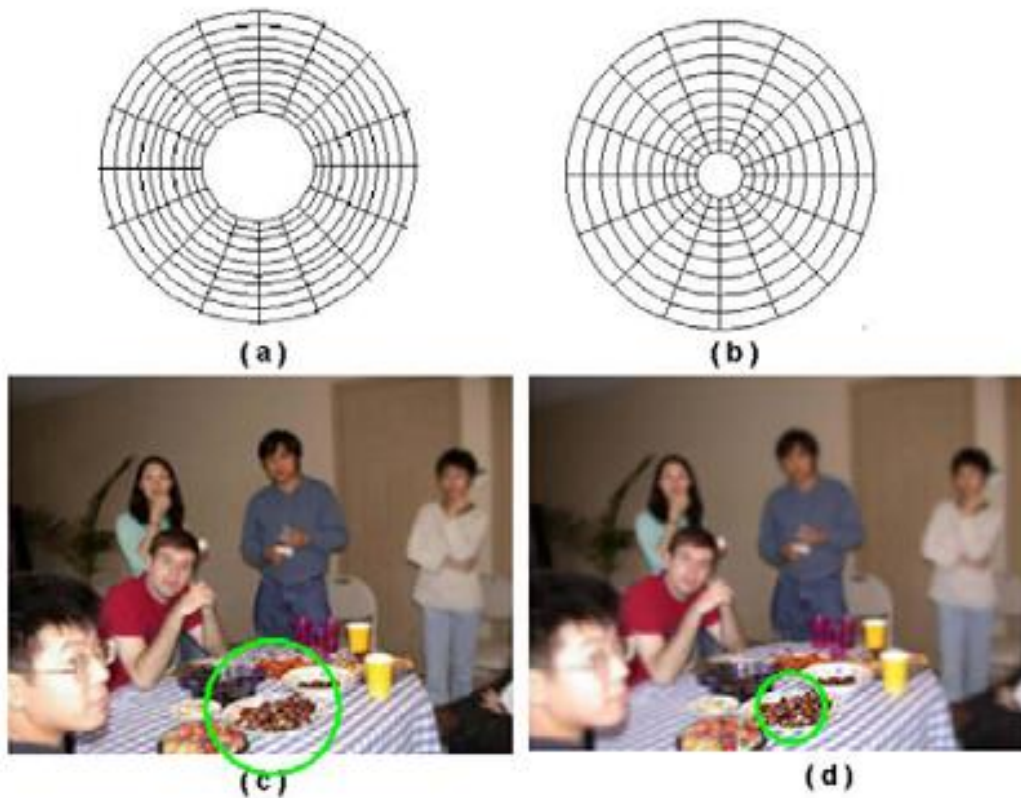
Uma vantagem importante dos modelos bayesianos é sua habilidade de aprender e de unificar muitos fatores em um fator principal. Pode-se dizer que tomam vantagem de estatísticas de cenas naturais ou outras características que atraem atenção.

Modelos de decisão teórica

Para esses modelos, o foco está em que a atenção visual deve ser conduzida levando-se em conta o objetivo final.

Um mapa de ativação é calculado da primeira vez por extração de características visuais primárias e por detectar objetos significativos da cena. Um filtro de retina adaptável como o usado por Gao et al. [32] é aplicado para esse mapa então gerar regiões de interesse ROI (*Region Of Interest*, ROI) cujas localizações correspondem a esses picos de ativação e cujos tamanhos são estimados por um algoritmo de ajuste iterativo. O foco da atenção é movido em série ao longo das ROIs detectadas por um mecanismo de decisão teórica. A sequência gerada de fixações oculares é determinada a partir de uma função de benefício perceptual com base nos custos de percepção e recompensas, enquanto a distribuição do tempo de diferentes ROIs é estimada pela aprendizagem e memória em decomposição.

A Figura 6 mostra um exemplo do filtro de retina adaptável. A semelhança com o olho humano determina a região de maior relevância em função da dilatação da retina.



(a-b) Partição do campo receptivo com campo de fixação para longe e pupila pequena.
(c-d) Resultado do filtro de retina correspondente (a-b).

Figura 6. Filtro de retina adaptável.

O filtro de retina adaptável é calculado por uma função logarítmica complexa com excentricidade. Trata-se de uma transformação logarítmica polar chamada *logmap* que foi estudada como uma boa aproximação ao seguimento retino cortical do sistema visual humano [40].

Modelos de decisão teóricas têm sido muito bem sucedidos em aplicações de visão computacional, tais como a classificação ao conseguir uma alta precisão na previsão de fixação.

Modelos de informação teórica

Esses modelos baseiam-se na premissa de que a computação localizada de relevância serve para maximizar a informação de uma amostra do ambiente. Eles lidam com a seleção das partes mais informativas de uma cena e descartam as partes restantes. Alguns autores desse modelo usam entropia condicional como Li et al.[33] Ele aproxima a entropia condicional pelo comprimento de codificação com perdas de dados multivariados de Gauss. O mapa final de relevância é acumulado por *pixels* e segmentado para detectar os objetos. Yan et al. [34] propõem uma nova versão desse modelo adicionando um esquema de multi-resolução.

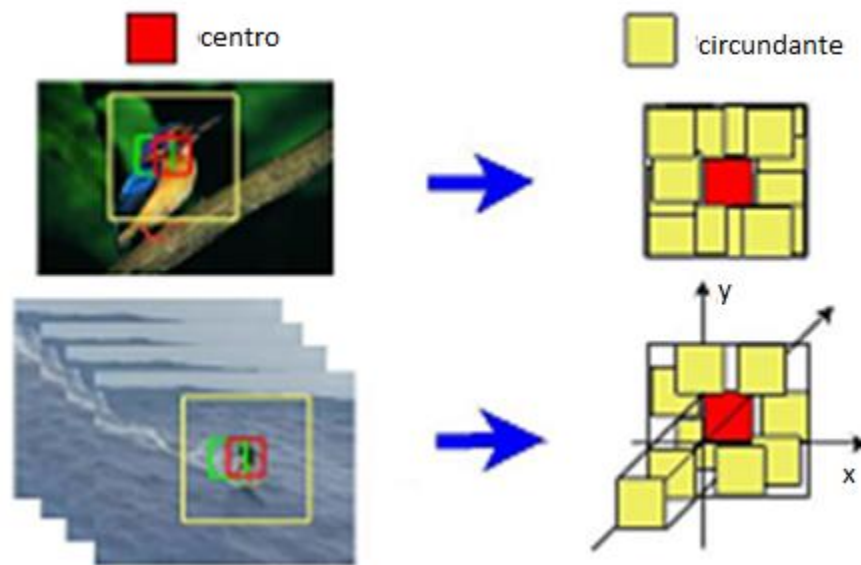


Figura 7. Modelo de informação teórica baseado na entropia.

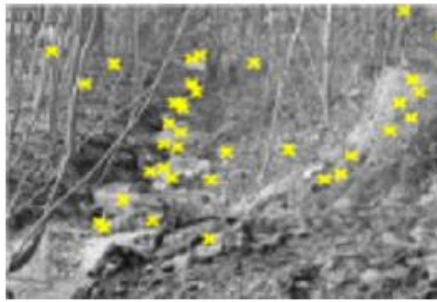
A Figura 7 mostra um exemplo desse modelo usando uma entropia condicional baseada no sistema visual humano. O quadrado do centro é a parte com maior relevância e os quadrados circundantes são regiões próximas ao centro.

Modelos gráficos

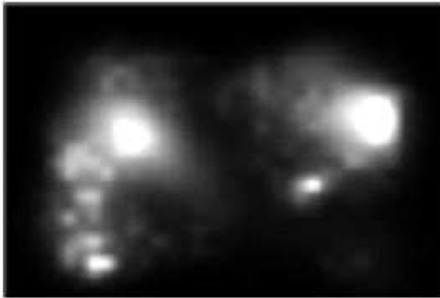
Um modelo gráfico consiste de uma estrutura probabilística em que um gráfico mostra a estrutura de independência condicional entre as variáveis aleatórias. Modelos de atenção nessa categoria tratam os movimentos dos olhos como em uma série temporal. Uma vez que existem variáveis ocultas que influenciam a geração de movimentos oculares, abordagens como o modelo oculto de Markov (*Hidden Markov Models, HMM*), redes bayesianas dinâmicas (*Dynamic Bayesian Networks, DBN*), e campo condicional aleatório (*Conditional Random Fields, CRF*) foram incorporadas.

Os modelos gráficos podem ser vistos como uma versão generalizada do modelo bayesiano. Assim tal visão permite uma modelagem de mecanismos de atenção mais complexa no espaço e no tempo, o que resulta em um maior poder de previsão.

As desvantagens residem na complexidade do modelo, especialmente quando se trata de treinamento e legibilidade, porque além de usar dados probabilísticos esses modelos usam outras camadas para gerar mais dados. A Figura 8 mostra um exemplo do uso desse método usado por Harel et al. [35] apresentam um modelo de relevância visual baseada em grafos (*Graph-Based Visual Saliency, GBVS*). Nesse modelo, extraem-se mapas de características em múltiplas escalas espaciais. Primeiro, uma pirâmide de escala e espaço é obtida das características da imagem: intensidade, cor e orientação (um trabalho semelhante ao do Itti et al [25]). Em seguida, um gráfico totalmente conectado sobre todos os locais de grade de cada mapa característico é construído. Pesos entre dois nós são atribuídos de forma proporcional à similaridade com os valores de recursos e sua distância espacial.



a. Imagem de mostra com fixação de olhos.



b. Mapa de relevância baseada em grafos.



c. Mapa de relevância tradicional.

Figura 8. Exemplo de modelo gráfico usado em [35].

Modelos de análise espectral

Nesse modelo ao invés de se processar uma imagem no domínio espacial, os *pixels* estão localizados no espaço da imagem e são determinados por suas coordenadas cartesianas (x , y) que correspondem à largura e ao comprimento, respectivamente. Esse procedimento deriva a relevância no domínio da frequência.

Usando a transformada de Fourier inversa, constrói-se um mapa de relevância no domínio espacial. O valor de cada ponto no mapa de relevância é então, guardado para indicar o erro de estimativa. Finalmente, o mapa de relevância é suavizado através de um filtro gaussiano para se conseguir um melhor efeito visual.

A Figura 9 mostra um exemplo desse modelo. A parte superior mostra uma imagem no domínio da frequência e parte inferior mostra uma imagem no domínio espacial.

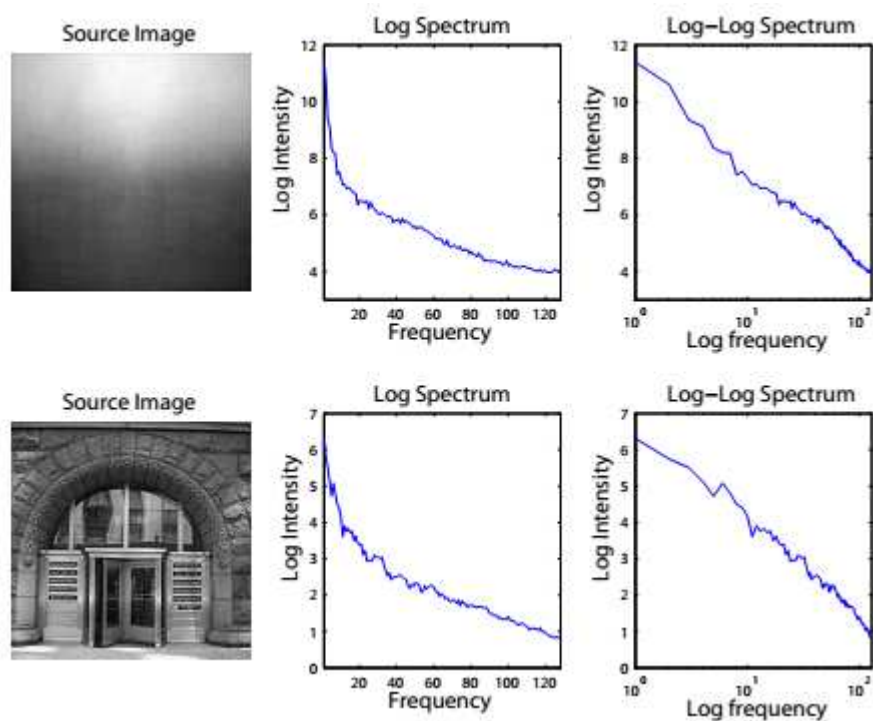


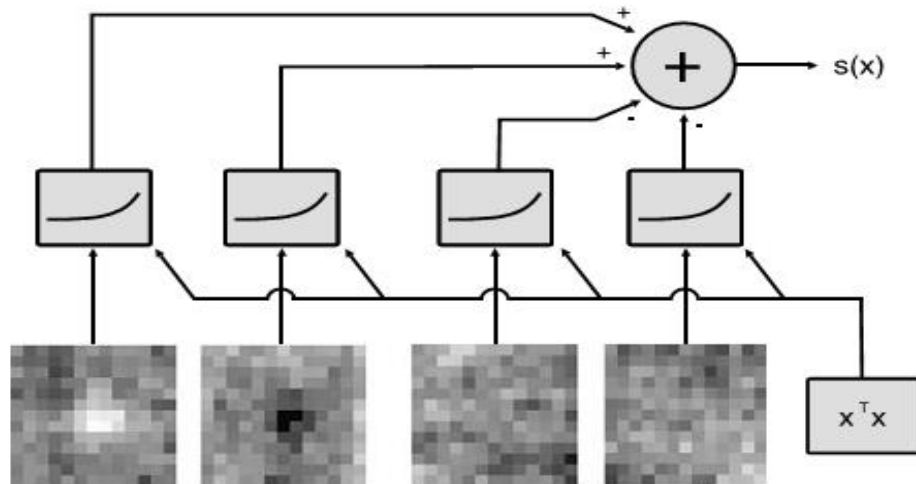
Figura 9. Exemplo de modelo usando a transformada de Fourier usado em [31].

Modelos de análise espectral são simples de se explicar e programar. Embora esses modelos tenham sido muito bem sucedidos, a semelhança com a biologia humana desses modelos ainda não é muito clara.

Modelos de classificação de padrões

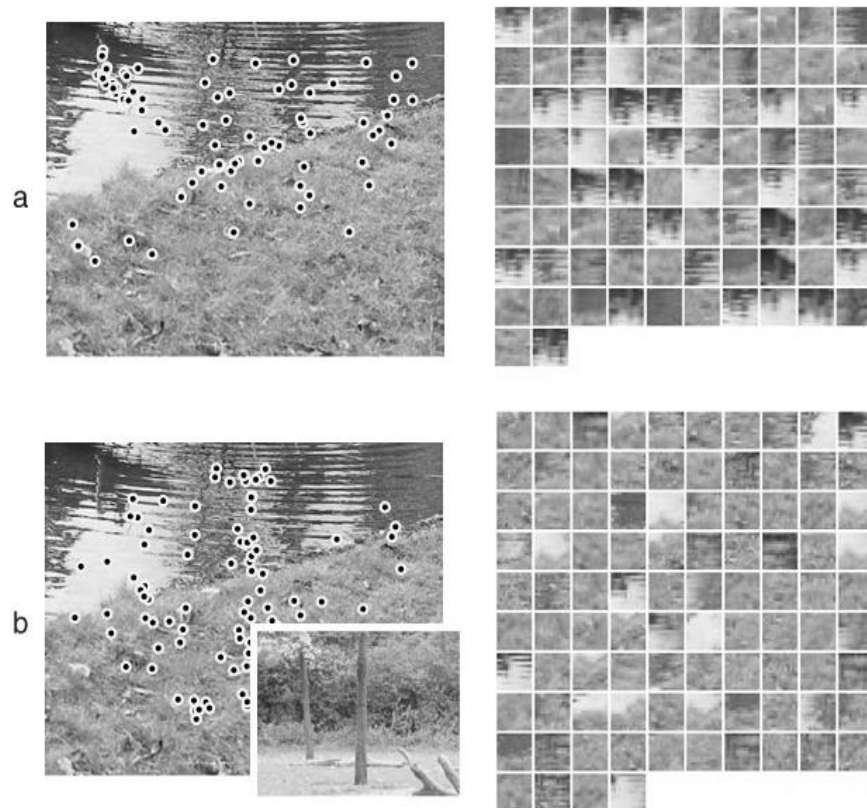
Abordagens de aprendizado de máquina também foram utilizadas na modelagem de atenção visual por modelos de aprendizagem a partir de dados gravados.

O modelo consiste em um mapeamento não linear a partir de um fragmento da imagem para um valor real, com treinamento para produzir resultados positivos sobre as fixações e saídas negativas sobre fragmentos de imagem selecionados aleatoriamente. A função de relevância é determinada pela sua maximização levando-se em conta um desempenho de previsão de dados observados. Uma máquina de vetor de suporte (*Support Vector Machine, SVM*) é treinada para determinar a relevância usando-se as intensidades locais como em Kienzle et al [36], que propõem apreender um conjunto de filtros temporais de fixação de olhos (*eye-fixation*) para se encontrar os locais interessantes. A vantagem dessa abordagem é que ela não precisa de suposições a priori sobre os recursos que contribuem para a relevância ou como esses recursos são combinados em um único mapa de relevância. Além disso, esse método produz operadores de centro-periferia análogos aos campos receptivos de neurônios em áreas visuais.



$S(x)$ é o valor de relevância e $x^T x$ é o sinal

Figura 10. Exemplo de padrão usado em [36].



- a. Pontos de fixação selecionados de fragmentos em base de dados.
- b. Pontos de fixação extraídos usando a imagem pequena.

Figura 11. Exemplo de classificação de padrões [36].

As Figura 10 e a Figura 11 mostram um modelo de padrões e sua aplicação em imagens, destacando-se os pontos de relevância. A Figura 12 mostra mapas de relevância de diferentes modelos estudados em [15]. Apresentam-se dessa forma, as semelhanças visuais dos resultados obtidos para cada modelo. Esses mapas são relevantes nesse trabalho para se definir as etapas do algoritmo.

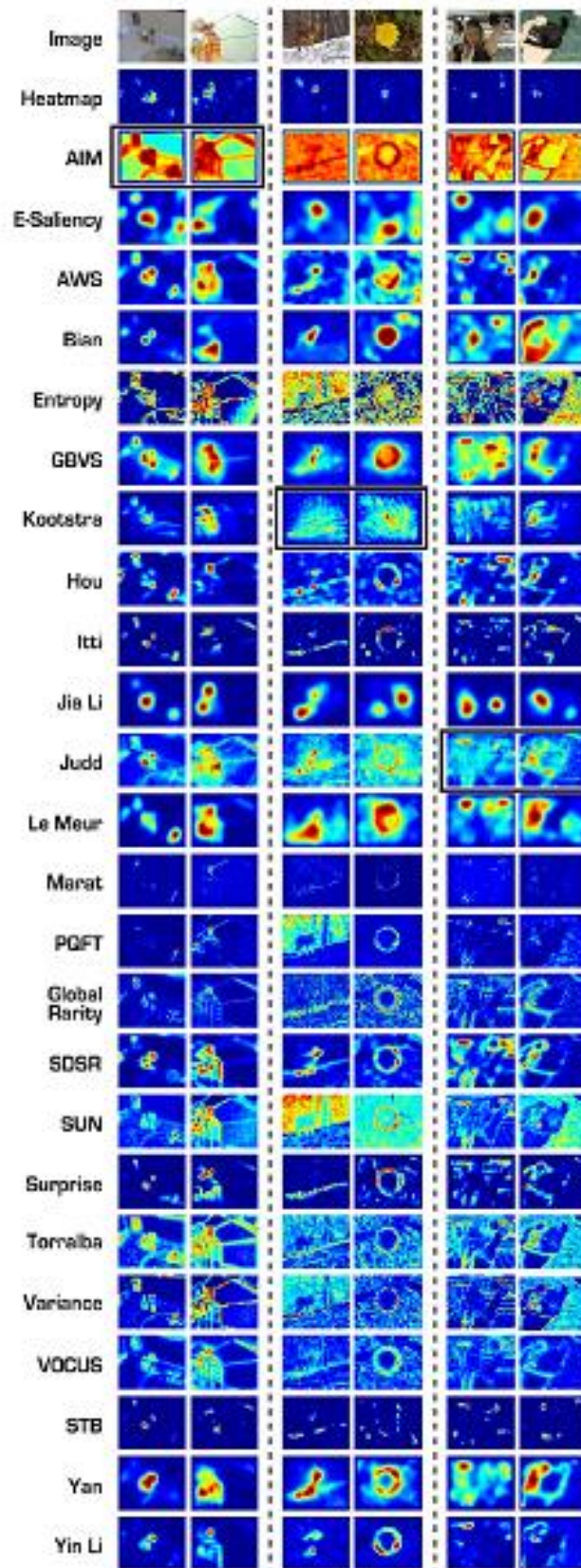


Figura 12. Exemplo de mapas de relevância dos modelos estudados em [15].

1.5 Relevância dos modelos estudados

Os modelos revisados anteriormente formam uma base para o desenvolvimento deste trabalho. Tem-se centrado no componente *bottom-up* de atenção visual. Embora todos os esforços anteriores sejam apreciados, o campo da atenção visual ainda carece de princípios computacionais, para a atenção dirigida à tarefa [30].

Por isso apresenta-se uma contribuição neste presente trabalho adicionando-se uma modificação no algoritmo de atenção visual proposto por L. Q. Chen et al. em [19]. Além disso, neste trabalho utiliza-se um banco de dados que proporciona um melhor casamento entre as tarefas executadas por pessoas e aquelas automatizadas pelo algoritmo. As medidas objetivas que foram implementadas mostram que com essas contribuições há uma redução da desvantagem existente nos enfoques atuais usando-se apenas avaliações subjetivas.

1.6 Organização da tese

Este trabalho está assim organizado: no capítulo 2 apresentam-se as bases do modelo proposto e alguns resultados preliminares que permitiram chegar ao algoritmo final; no capítulo 3 descrevem-se a implementação, experimentos e algoritmos usados; no capítulo 4 mostram-se os resultados obtidos como prova de conceito do modelo de atenção visual usado já com as contribuições deste trabalho; e no capítulo 5 apresentam-se as conclusões e são propostos os possíveis trabalhos para o futuro.

CAPÍTULO 2 - DESCRIÇÃO DO MODELO PROPOSTO

Neste capítulo se descreve brevemente um modelo de atenção visual proposto.

2.1 Descrição do modelo de atenção visual

Para uma escolha do modelo de atenção usou-se como ponto de partida o trabalho [13] porque se trata de um dos modelos que estão no estado da arte e serve aos propósitos deste trabalho. O modelo de atenção visual define uma imagem como um conjunto de objetos de atenção. Inicia-se com a seleção da imagem. Em seguida, ela é analisada por três algoritmos de detecção (relevância, faces e texto). Cada um deles determina um mapa de atenção os quais são combinados em um só mapa. Nesse mapa definem-se os objetos de atenção. As combinações de objetos de atenção servem para se encontrar as possíveis soluções que se aproximam ao tamanho alvo, concluindo assim a adaptação da imagem. Esse processo é mostrado na Figura 13.

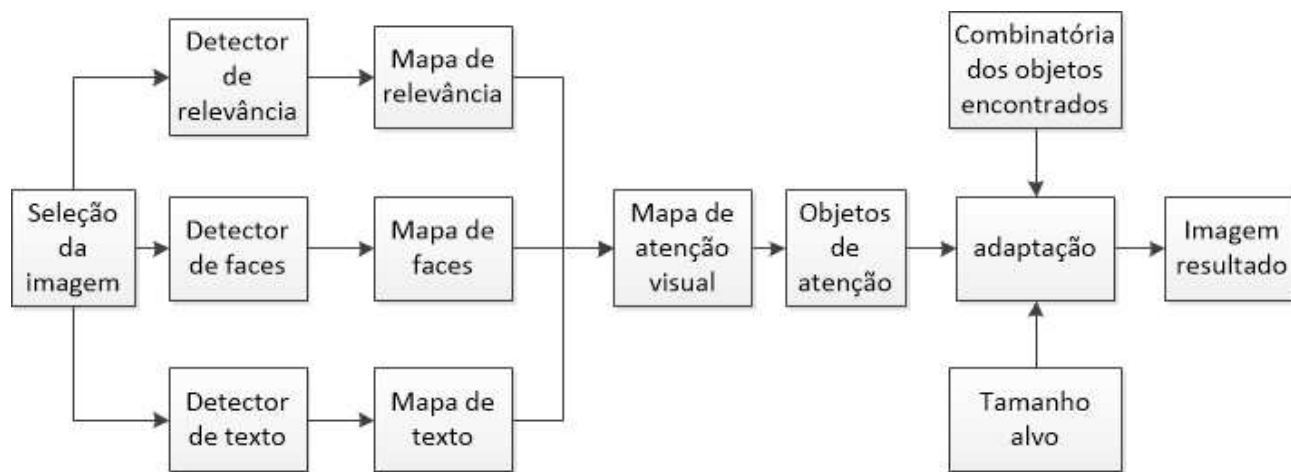


Figura 13. Diagrama de fluxo do modelo proposto.

Um objeto de atenção (*Attention Object*, AO) é uma informação considerada como foco de atenção. Assim, em geral um AO representa um objeto semântico onde é possível de reconhecer, por exemplo, a face, texto, cor, etc.

O modelo de atenção aplicado em uma imagem gera como resultado um conjunto de objetos de atenção. A Figura 3 mostra os objetos encontrados de três tipos: em azul os objetos de tipo texto, em vermelho o objeto de relevância e em verde os objetos de tipo face. Para cada AO se quantiza um valor da atenção. Definem-se três atributos a cada objeto de atenção: regiões de interesse (*Region Of Interest*, ROI), valor de atenção (*Attention Value*, AV) e o mínimo tamanho perceptível (*Minimum Perceptible Size*, MPS).



Figura 14. Objetos de atenção em uma cena. A imagem foi tomada de um banco de dados [44].

Região de interesse (ROI): Existem partes em uma imagem que podem ser definidas como mais importantes que outras. Essa característica permite ao usuário, por exemplo, definir que parte da imagem é codificada com melhor qualidade e menor distorção do que o restante da imagem [23]. A região de interesse é utilizada para determinar a região do objeto de atenção.

Valor de atenção (AV): Em uma cena pode-se ter mais de um objeto de atenção. Para distinção entre esses objetos deve-se ter uma medida que diferencie sua importância. Assim, um valor que quantifica a atenção é definido para cada um dos objetos de atenção. Define-se uma faixa de $0 \leq AV \leq 1$ para a aplicação desse valor.

Mínimo tamanho perceptível (MPS): Um objeto pode ser trivialmente um *pixel* e isso é considerado incoerente. Para evitar no modelo este tipo de incoerência, é definido um tamanho mínimo para um objeto de atenção de acordo com o tamanho da cena total no qual seja possível perceber o objeto. Como é mostrado em [24], é adequado considerar que a mínima resolução necessária. Assim, para os sistemas de reconhecimento de faces é de 24x24 conforme descrito para o algoritmo de Viola e Jones [27]. Esse tamanho é escolhido neste trabalho. No entanto, existem outras configurações como: 19x19 *pixels* usado por Sung e Poggio [37]; 20x20 por Rowley et al [38]; e 58x58 por Heisle et al. [39].

Cada objeto de atenção fica definido por:

$$\{AO_i\} = \{(ROI_i, AV_i, MPS_i)\}, 1 \leq i \leq N$$

(3)

Onde:

N é o número total de objetos de atenção encontrados na imagem.

AO é o objeto de atenção;

AV é o valor de atenção;

MPS é o mínimo tamanho perceptível.

2.2 Atenção visual seletiva

A atenção visual envolve diversos fatores. Uma forte evidência para a seleção espacial vem de estudos de sinais que mostram que a detecção de um objetivo é melhorada quando esta é precedida por um sinal válido [25]. Cada abordagem de atenção visual gera um mapa de atenção seletivo independente. Neste trabalho utilizam-se três tipos de atenção: relevância, texto e face.

2.2.1 Atenção visual utilizando relevância

O modelo de atenção visual utilizado neste trabalho é baseado na relevância e definido em [17]. A entrada visual é decomposta em um conjunto de mapas de características topográficas (cor, intensidade e orientação). A combinação de informação dos mapas individuais de características forma um mapa de relevância que é uma medida global da visibilidade na cena [11].

Para localizar o foco de atenção humana, serão detectadas as regiões que são mais atrativas à atenção humana. Inicialmente, transforma-se o mapa de relevância em binário (0, 1), onde o limiar é estimado de maneira adaptativa [18].

Como resultado define-se o valor da atenção visual de relevância como:

$$AV_{relevância} = \sum_{(i,j \in R)} B_{i,j} \cdot W_{relevância}^{i,j}$$

(4)

Onde:

$AV_{relevância}$ é o valor de atenção para a relevância;

$B_{i,j}$ é o nível em escala de cinza do *pixel* no ponto (i,j) na região de relevância R;

$W_{relevância}^{i,j}$ é o peso de posição desse *pixel*.

Em geral, uma pessoa enfoca mais atenção na região próxima ao centro da imagem. Assim, uma função gaussiana normalizada e vinculada à imagem é usada para se determinar o peso de posição. Por exemplo, na Figura 15 define-se uma função gaussiana para determinar a relevância da cor no sistema RGB (*Red, Green and Blue*). Na Figura 16 apresenta-se um exemplo da aplicação do algoritmo de relevância.

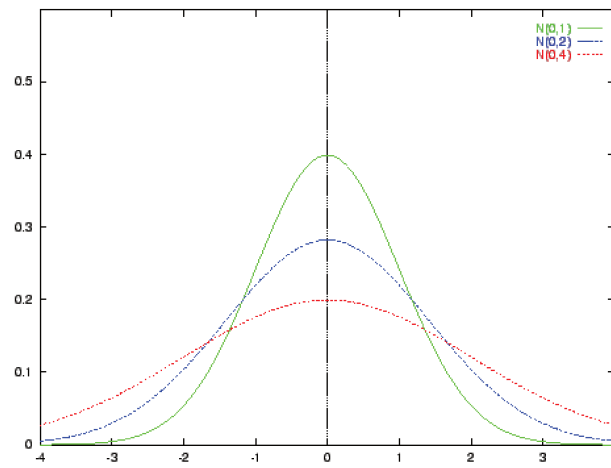
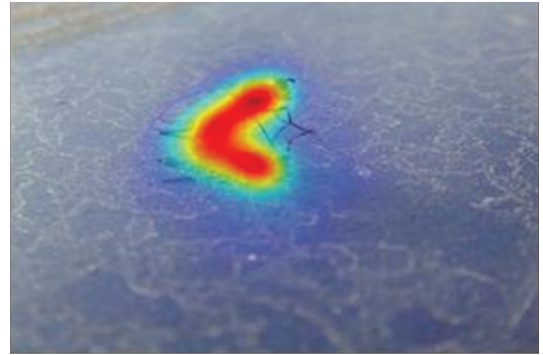


Figura 15. Função gaussiana para a relevância da cor.



a. Imagem de amostra para detecção da relevância.



b. Mapa de relevância da imagem.



c. Objeto de relevância.

Figura 16. Exemplo da atenção visual usando um algoritmo de relevância.

2.2.2 Atenção visual utilizando a face

As faces ou rostos são uma das características mais notórias na percepção visual humana. A presença de rostos na imagem gera regiões de atenção. Por esse motivo deve ser integrado em modelos de atenção visual.

O reconhecimento de uma face ou rosto (a partir de imagens fotográficas e imagens de vídeo) está emergindo como uma atividade na área de pesquisa com numerosas aplicações comerciais e por pressão de lei. Essas aplicações requerem algoritmos robustos para reconhecimento de faces humanas sobre diferentes condições de iluminação, expressões faciais e orientações. Um esquema geral usado para reconhecimento de face é ilustrado na Figura 17 [41].

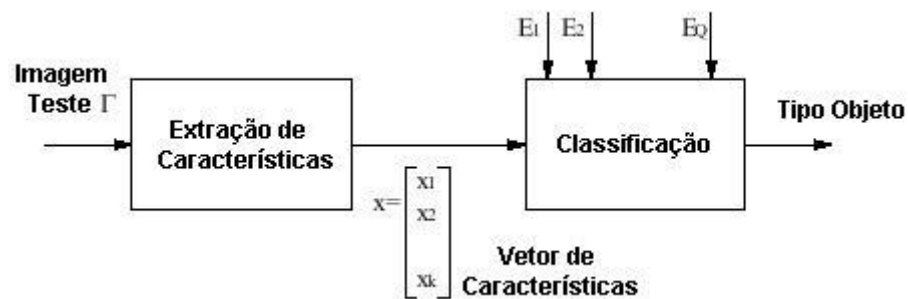


Figura 17. Esquema geral de reconhecimento de faces.

O vetor de características $x = [x_1, x_2, \dots, x_k]$ extraído de uma imagem teste é comparado a cada um dos vetores de características extraídos de todos os exemplos de imagens de face E_1, E_2, \dots, E_Q . Uma medida de similaridade no espaço de características é usada para classificar a imagem de entrada como uma das imagens de teste. A razão entre o número de imagens de face classificadas corretamente sobre o número total de faces classificadas pelo sistema de reconhecimento define o desempenho de reconhecimento (taxa de reconhecimento) do sistema. Baseado na extração de características e técnicas de classificação usadas, as abordagens de reconhecimento de face segundo [27] são: características, funções de classificação de aprendizagem e arquitetura de atenção em cascata.

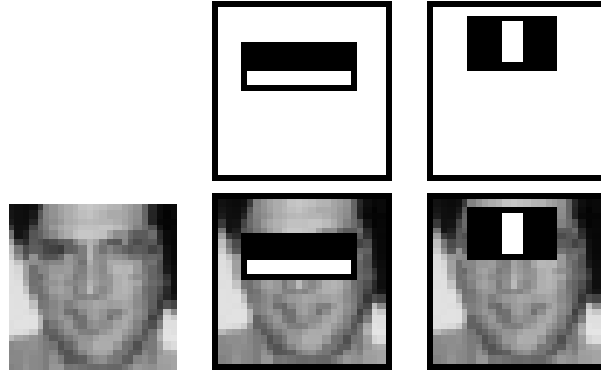


Figura 18. Exemplo de características básicas usadas em [19].

Faz-se uso do algoritmo descrito em [19] o qual combina um conjunto de características básicas como é mostrado na Figura 18, a fim de se formar uma cascata de classificadores que juntos resultam em um eficiente detector de objetos, nesse caso de faces.

Aplica-se o algoritmo e obtém-se uma informação das faces presentes na imagem. Conhecendo-se sua região e posição, define-se:

$$AV_{face} = \sqrt{Area_{face}} \times W_{face}^{pos} \tag{5}$$

Onde:

AV_{face} é o valor de atenção para o face;

$Area_{face}$ é o tamanho da região do rosto detectado;

W_{face}^{pos} é o peso de objeto face na posição “pos”.

Na Figura 19 mostra-se um exemplo do algoritmo aplicado em uma imagem que tem um rosto. A imagem usada é obtida de um banco de dados [44].



Figura 19. Exemplo de detecção de face.

2.2.3 Atenção visual utilizando um texto

Na percepção humana, a presença de texto em uma cena chama a atenção visual em muitas situações. Esses textos contêm informação semântica e são úteis para se definir um objeto de atenção.

O algoritmo de detecção de texto usado neste trabalho é baseado em [20] e provê uma região e uma posição com o qual se pode definir um valor de atenção como se segue:

$$AV_{texto} = \sqrt{Area_{texto}} \times W_{texto}$$

(6)

Onde:

AV_{texto} é o valor de atenção para o texto;

$Area_{texto}$ é o tamanho da região do texto encontrado;

W_{texto} é peso da região de texto encontrado.



Figura 20. Exemplo de detecção de texto.

2.2.4 Combinação dos mapas de atenção

As combinações de muitas medidas de atenção visual requerem o ajuste de cada uma delas para integração. Em alguns cenários pode ser que os rostos e texto sejam mais prioritários que a relevância. Assim, o valor de atenção é definido como se segue:

$$AV_i = W_k \cdot AV_i^k / \sum_i AV_i^k \quad (7)$$

Onde:

W_k é o peso do modelo k ;

AV_i^k é o valor de atenção do AO_i detectado no modelo k .

Para determinar o peso de atenção de cada modelo existe a dependência do tipo de problema e a aplicação onde será usado o mapa de atenção final.

Neste trabalho os pesos são definidos supondo-se uma aplicação para a publicidade, onde a maior prioridade ou peso é para a face, seguido pelo texto e pela relevância.

2.3 Adaptação de imagens

A apresentação de imagens depende das propriedades dos dispositivos de saída como a resolução e tamanho das telas. Esses fatores são importantes e devem ser levados em conta para se chegar a uma apropriada apresentação da imagem.

A adaptação de imagens consiste na apropriada apresentação de conteúdo visual de uma imagem em um conjunto de vistas e restrições dadas pelo usuário [3].

O problema de adaptação de imagem pode ser tratado como uma manipulação de objetos de atenção para prover tanta informação quanto for possível. No caso da adaptação da imagem em uma região de tamanho restrito, baseado no modelo descrito anteriormente, o problema se transforma em um problema de programação linear inteira. Isso porque se tem um conjunto de objetos de atenção, os quais se deseja separar formando então a região objetivo. As possíveis soluções são um conjunto de objetos com melhor ajuste à região objetivo. Devem-se prover alguns mecanismos descritos a seguir para se avaliar as opções de soluções.

2.3.1 Fidelidade de informação

Para avaliar de forma quantitativa a solução ótima toma-se um valor de fidelidade da informação (*Information Fidelity*, IF).

$$IF(R) = \sum_{ROI_i \subset R} AV_i \cdot IF_{AO_i}$$

(8)

Onde:

$IF(R)$ é fidelidade da informação de uma região R;

AV é valor de atenção;

IF_{AO_i} é fidelidade de informação para o objeto AO_i ;

ROI é a região de interesse.

2.3.2 Algoritmo de adaptação

Com o uso do modelo escolhido, o problema de adaptação converte-se na procura do melhor uso da área objetivo definida como T que represente à imagem original. Esquemas diversos podem ser aplicados para se encontrar diferentes resultados. Cada resultado corresponde a uma única solução que pode ser representada pela região R na imagem original.

Para propósitos deste trabalho tem-se que a área objetivo é menor do que a região original. Uma medida objetiva para a fidelidade de informação de uma imagem adaptada pode ser formulada como segue:

$$\begin{aligned} IF(R) &= \sum_{ROI_i \subset R} AV_i \cdot IF_{AO_i} \\ &= \sum_{ROI_i \subset R} AA_i \cdot u(r_R^2 \cdot size(ROI_i) - MPS_i) \end{aligned} \tag{9}$$

Onde:

$u(x)$ é definida como:

$$u(x) = \begin{cases} 1 & \text{se } x \geq 0 \\ 0 & \text{em outro caso} \end{cases}$$

$size(x)$ é uma função que calcula a área de uma ROI.

r_R denota a taxa de redimensionamento da imagem em área retangular que pode ser calculada com:

$$r_R = \min\left(\frac{Width_T}{Width_R}, \frac{Height_T}{Height_R}\right) \quad (10)$$

Onde:

Width é a largura da área retangular;

Height é a altura da área retangular.

As medidas da área objetivo T e da região solução R, transformam o problema de fazer uma decisão de adaptação em o problema de se procurar uma região que contenha um conjunto ótimo de objetos de atenção.

$$\begin{aligned} \max(IF(R)) = \\ \max_R \left\{ \sum_{ROI_i \subset R} AV \cdot u(r_R^2 \cdot size(IR_i) - MPS_i) \right\} \end{aligned} \quad (11)$$

Onde:

IF(R) é fidelidade da informação de uma região R;

AV é valor de atenção;

size(x) é uma função que calcula a área de uma ROI;

MPS é o mínimo tamanho perceptível.

Para uma imagem de largura *m* e altura *n*, a complexidade para encontrar uma solução ótima é de $O(m^2n^2)$. Se o tamanho de *m* ou *n* for muito grande, então o custo computacional pode ser muito alto. No entanto, com a fidelidade de informação de uma região adaptada pode-se reduzir o tempo computacional procurando-se um conjunto ótimo de objetos de atenção antes de se gerar a solução final.

2.3.3 Conjunto de objetos de atenção válidos

Define-se I como um conjunto de objetos de atenção

$$I \subset \{AO_1, AO_2, \dots, AO_N\}$$

Deve-se encontrar um conjunto de objetos de atenção que tenha a mais alta fidelidade de informação depois da adaptação. Um conjunto de objetos de atenção é válido se

$$\frac{MPS_i}{size(ROI_i)} \leq r_I^2, \forall AO_i \in I \quad (12)$$

Onde R_I é a taxa de redimensionamento de R_I para T que é calculado como se segue:

$$r_I = \min\left(\frac{Width_T}{Width_I}, \frac{Height_T}{Height_I}\right) \quad (13)$$

Onde:

$$Width_I = \max_{AO_i, AO_j \in I} |Right_i - Left_j|$$

e

$$Height_I = \max_{AO_i, AO_j \in I} |Bottom_i - Top_j|$$

Notemos que os valores $Width_I$ e $Height_I$ definem a região R_I , e $Right_i, Left_j, Bottom_i, Top_j$ são os limiares do objeto de atenção. Para validar um objeto de atenção usa-se uma taxa de redimensionamento, para quaisquer dois conjuntos de objetos de atenção I_1 e I_2 se $I_1 \subset I_2$ e I_1 é inválido então I_2 é inválido.

2.3.4 Conjunto ótimo de objetos de atenção

A procura de um conjunto ótimo é um problema de programação linear que pode ser resolvido com um algoritmo de *branch and bound*, que serve para encontrar soluções ótimas em problemas de otimização. Consiste em uma enumeração sistemática de todos os candidatos à solução, em que grandes subconjuntos de candidatos infrutíferos são descartados em massa utilizando-se os limites superior e inferior da quantidade ótima.

Para programar esse algoritmo é usada uma árvore binária na qual cada nó denota um conjunto diferente de objetos de atenção. Cada nível corresponde à inclusão de um objeto, e cada bifurcação define se o objeto será mantido ou não. A altura da árvore é o número de objetos de atenção dentro da imagem e cada folha corresponde a um possível conjunto de objetos. Em cada inserção é verificado se um conjunto de objetos é válido e se atualiza a proximidade com relação à fidelidade de informação. Se isso não ocorrer, o nó é truncado e a complexidade computacional é reduzida.

2.3.5 Adaptação da solução na região objetivo

Após encontrar uma solução ótima deve-se realizar uma adaptação de acordo com as restrições da região objetivo. Para tanto, pode ser escalado o conjunto I dentro da região objetivo mantendo a proporção de modo que os objetos de atenção fiquem na região objetivo.

A taxa de escalamento deve ser:

$$\tau_I^{max} = \max_{AO_i \in I} (MPS_i / size(ROI_i))$$

(14)

Onde:

size(x) é uma função que calcula a área de uma ROI;

MPS é o mínimo tamanho perceptível.

Na etapa de escalamento é necessário definir um limiar de adaptação que maximize o uso da região objetivo que é definida pela quantidade da área restante da região.

$$L_I = Area_{objetivo} - Area_I$$

(15)

Onde:

L_I é o limiar de adaptação do conjunto I;

$Area_{objetivo}$ é a área objetivo;

$Area_I$ é a área do conjunto I.

Algumas possíveis soluções para o exemplo da Figura 14 são mostradas na Figura 21.

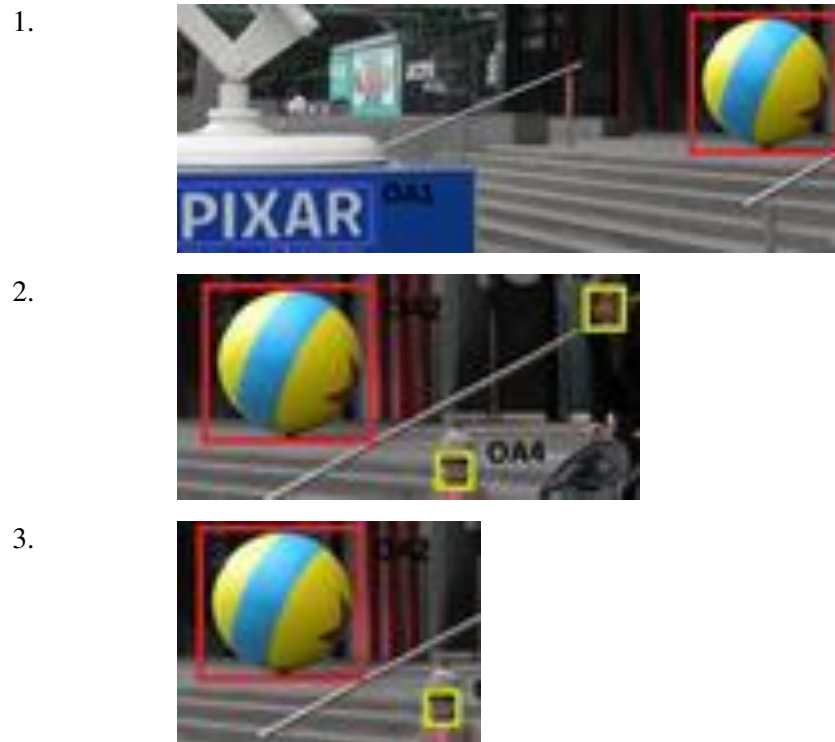


Figura 21. Possível conjunto de 3 objetos de atenção.

CAPÍTULO 3 - METODOLOGIA

Para realizar a prova de conceito do modelo de atenção visual utilizado neste trabalho já com as inserções propostas faz-se uma comparação de dados baseados em 3 especialistas usando-se o mesmo banco de dados de [44] com o modelo proposto e o modelo original. As imagens possuem variedade de categorias, incluindo animais, pessoas, prédios e diversidade de objetos. Cada imagem é recortada manualmente pelos especialistas para então se realizar uma comparação com os resultados obtidos automaticamente pelo modelo proposto neste trabalho.

A comparação é feita com a cálculo de duas distâncias. A distância euclidiana para os centróides e a distância de Mahalanobi para os centróides e regiões. Obtém-se dos grupos de dados os resultados, que serão comparados com um cálculo da distância de Mahalanobi usando-se um princípio de construção de contraste de hipótese. Esse princípio de junção-interseção é um princípio geral que permite construir contrastes multivariados a partir de contrastes univariados. Inicialmente usa-se a hipótese nula $H: BD = MP$. Nessa hipótese, BD constitui os dados do banco de dados e MP constitui os dados obtidos pelo modelo proposto neste trabalho.

3.1.1 Distância de Mahalanobis

Em estatística, a distância de Mahalanobis é uma medida natural da diferença entre as médias de dois grupos levando-se em conta as covariâncias. Está baseada nas correlações entre variáveis com as quais distintos padrões podem ser identificados e analisados. É uma estatística útil para determinar a similaridade entre uma amostra desconhecida e outra conhecida. Essa distância distingue-se da distância euclidiana já que ela leva em conta as correlações do conjunto de dados e é invariante com a escala, ou seja, não depende da escala das medições [42].

Formalmente, a distância de Mahalanobis entre um grupo de valores com média $\mu = (\mu_1, \mu_2, \mu_3, \dots, \mu_p)^T$ e matriz de covariância S para um vetor multivariável $x = (x_1, x_2, x_3, \dots, x_p)^T$ é definida como:

$$D_M(x) = \sqrt{(x - \mu)^T S^{-1} (x - \mu)}$$

(16)

Onde:

D_M é a distância de Mahalanobis;

x é o vetor multivariável;

μ é media de valores.

CAPÍTULO 4 - RESULTADOS

Neste capítulo apresentam-se os resultados obtidos pelo modelo de atenção visual proposto em comparação com os resultados obtidos manualmente do banco de dados. Esse experimento foi realizado sobre um conjunto de 235 imagens com conteúdos diferentes ou variados.

Para o algoritmo de adaptação necessita-se de um ponto que represente a maior atenção da imagem. A partir desse ponto, deve-se fazer um corte na imagem adaptado ao objetivo alvo. Do banco de dados tem-se três (são 3 especialistas) áreas selecionadas como de maior importância. Calcula-se seu centróide e então se faz uma comparação com aqueles obtidos pelo algoritmo proposto. A Figura 14 mostram os resultados obtidos. Nessa figura, têm-se os valores de posições dos centróides encontrados.

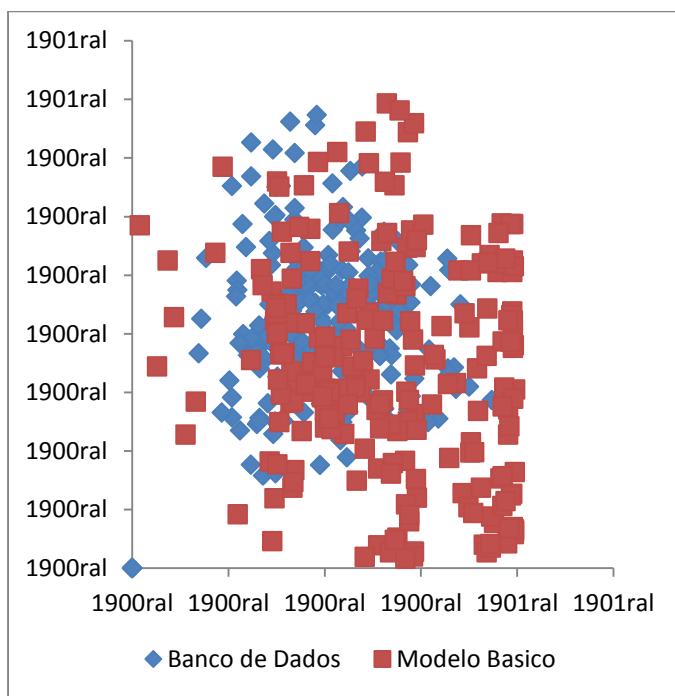


Figura 22. Gráfico de dispersão dos resultados para o modelo base original e para o banco de dados. Valores médios.

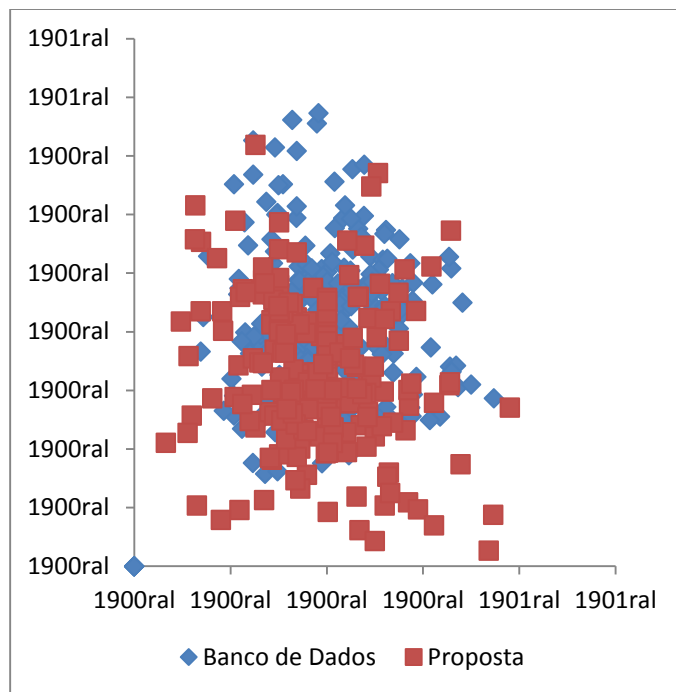


Figura 23 Gráfico de dispersão dos resultados para o sistema proposto e para o banco de dados.

A matriz de covariância calculada é:

$$\begin{pmatrix} 3905,88 & -50,79 \\ -50,79 & 4001,16 \end{pmatrix}$$

A distância de Mahalanobi calculada dos conjuntos de dados é de 0,71. Esse resultado é menor do que "1". Isso significa que a similaridade encontrada é menor do que um *pixel* mostrando assim que o algoritmo se aproxima muito bem dos dados de teste do banco de dados.

Distância Modelo	Distância Euclidiana	Distância de Mahalanobi
Modelo base	140,77	0,89
Modelo Proposto	107,92	0,71

Tabela 1. Comparação das distâncias calculadas.

No Apêndice 1 apresentam-se os fluxogramas dos programas referentes aos modelos base e proposto.

No Apêndice 2 apresentam-se partes dos códigos de programas pertinentes.

No Apêndice 3 mostram-se os resultados tabelados usando-se um banco de dados, cujo link é fornecido neste trabalho.

Link para o banco de dados [44].

http://research.microsoft.com/en-us/um/people/jiansun/SalientObject/salient_object.htm

No Apêndice 4 mostram-se algumas figuras pertinentes.

CAPÍTULO 5 - CONCLUSÕES E TRABALHO FUTURO

Neste trabalho foi apresentado um modelo de atenção visual para o problema de adaptação de imagens em regiões de diferentes tamanhos. Tomou-se como ponto de partida um modelo base original que é modificado para se obter um melhor desempenho em relação à uma medida denominada distância de Mahalanobi.

Para se compor a presente proposta foram aproveitados os mecanismos de percepção humana. A utilização de tais mecanismos permite definir um modelo de atenção visual.

No capítulo 1 foram estudados os diferentes modelos de atenção visual encontrados na literatura. Um enfoque maior é dado ao modelo encontrado em [19] por ser um modelo básico. Nesse modelo, a adaptação de imagens está focada em se encontrar objetos de atenção baseando-se em três algoritmos que determinam um valor de importância na imagem tais como relevância, texto e face. Esses parâmetros determinam uma região de maior importância na imagem. Uma vez determinada essa região é possível proceder com o uso de um algoritmo de adaptação. Esse algoritmo usa um método de programação linear o qual acopla os objetos encontrados previamente à uma nova região determinada pelo usuário para maximizar seu valor de atenção. Os resultados experimentais referentes à esse modelo mostram que ocorrem casos em que objetos encontrados pelos 3 diferentes algoritmos coincidem. Nessa situação, por exemplo, um objeto de relevância pode ser associado a um rosto e ao mesmo tempo é considerado como outro objeto pelo algoritmo de detecção de faces. Obtém-se assim, dois objetos considerados distintos, embora eles estejam na mesma área, ou seja, na mesma área de localização. Na presente proposta isso não ocorre. De fato, evita-se a concorrência de objetos por encontrar primeiro mapas de atenção (relevância, face ou texto). Em seguida, esses mapas são combinados em um só mapa de atenção. Somente depois de se obter esse mapa de atenção visual dá-se continuidade ao restante do processo com a segmentação do mapa em objetos de atenção. Em seguida, continua-se com o algoritmo de adaptação.

Além disso, para melhorar a velocidade de cálculo do algoritmo a utilização de processadores gráficos (*Graphics Process Unit, GPU*) é mais adequada porque os algoritmos utilizados no modelo de atenção tem um alto grau de paralelização.

5.1 Análise dos resultados e comentários gerais

Dos resultados obtidos pode-se concluir que a proposta tem 23% de melhor aproximação pela distância Euclidiana e de 20% pela distância de Mahalanobi comparando-se com os resultados originais obtidos com os dados do banco de dados. Isso confirma que o uso de modelos de atenção visual melhora o desempenho dos algoritmos de adaptação.

Utilizaram-se os dois enfoques dos modelos de atenção visual: *top-down* (faces e texto) e *bottom-up* (relevância) combinando-os a fim de se conseguir uma melhor proposta de solução.

Como passo futuro deve-se realizar uma avaliação com uma base de imagens mais extensa. Além disso, podem-se obter eventuais melhorias dos modelos de atenção visual para aplicações específicas como *marketing* visual e detecção automática de regiões de interesse.

5.2 Contribuições

Em primeiro lugar, utiliza-se o algoritmo de adaptação proposto para melhorar as apresentações de imagens de grande tamanho em telas de menor capacidade. Esse algoritmo também pode ser adaptado para uso em navegação por visão artificial já que enfoca a imagem em certo ponto de direção.

Por último propõe-se uma medida para uma comparação de algoritmos de adaptação de forma objetiva evitando-se assim perguntas subjetivas para a avaliação de um método em estudo.

REFERÊNCIAS

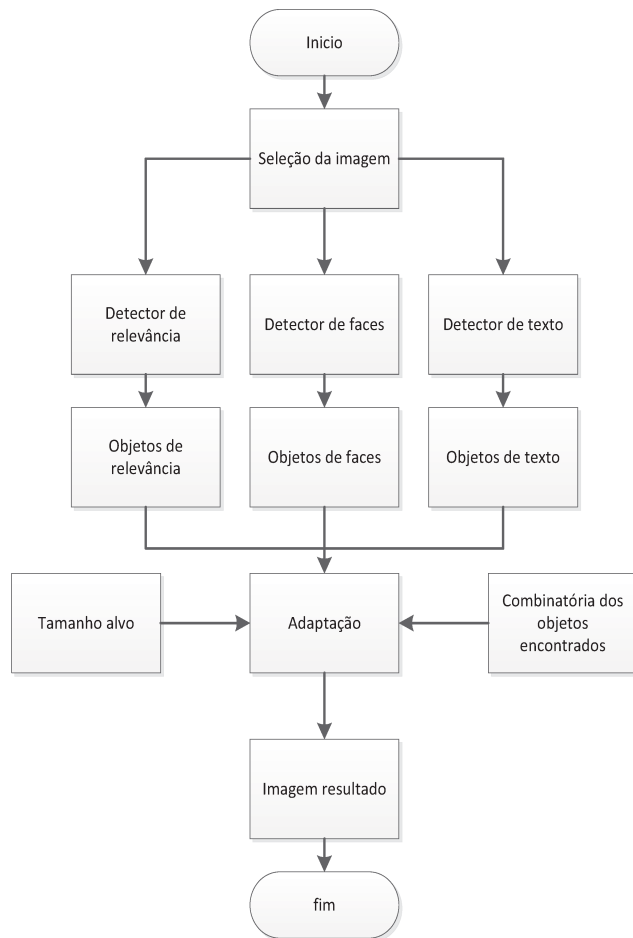
- [1] J. Yan, S. Lin, and S. B. Kang, “*Learning the Change for Automatic Image Cropping*”. pp. 971–978, 2013.
- [2] C. Wang, Y. Liang, D. Wang, X. Luo, and Z. Su, “*Optimised image retargeting using aesthetic-based cropping and scaling*”. *IET Image Processing*, vol. 7, no. 1, pp. 61–69, Feb. 2013.
- [3] R. Rosenbaum and B. Hamann, “*Raster image adaptation for mobile devices using profiles,*”. p. 83040H–83040H–10, Feb. 2012.
- [4] Q. Zhang, H. Xiao, and H. Liu, “*A Hierarchical Computational Model of Selective Visual Attention*”. *Journal of Software*, vol. 6, no. 9, pp. 1705–1712, Aug. 2011.
- [5] W. Yin, J. Luo, and C. W. Chen, “*Semantic adaptation of consumer photo for mobile device access*”. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, pp. 1173–1176, May 2010.
- [6] M. A. Hasan and C. Kim, “*An Automatic Image Browsing Technique for Small Display Users*”. pp. 2044–2049, 2009.
- [7] T. Toriu, S. Nakajima, and H. Hama, “*Hierarchical Attention Focusing and a Display Method for Small Mobile Device*”. *2008 3rd International Conference on Innovative Computing Information and Control*, pp. 56–56, 2008.
- [8] G. Ciocca, C. Cusano, F. Gasparini, and R. Schettini, “*Self-Adaptive Image Cropping for Small Displays*”. *IEEE Transactions on Consumer Electronics*, vol. 53, no. 4, pp. 1622–1627, Nov. 2007.
- [9] X. Xie, H. Liu, and H. Zhang, “*Browsing large pictures under limited display sizes*”. *IEEE Transactions on Multimedia*, vol. 8, no. 4, pp. 707–715, Aug. 2006.

- [10] H. Liu, X. Xie, W.-Y. Ma, and H.-J. Zhang, “Automatic browsing of large pictures on mobile devices”. *Proceedings of the eleventh ACM international conference on Multimedia - multimedia '03*, p. 148, 2003.
- [11] J. E. Capó-aponte, L. A. Temme, H. L. Task, A. R. Pinkus, M. E. Kalich, A. J. Pantle, and C. E. Rash, “Visual perception and cognitive performance”. 1999.
- [12] J. M. Wolfe, “Visual Attention”. pp. 335–386, 2000.
- [13] K. Koch, J. McLean, R. Segev, M. a Freed, M. J. Berry, V. Balasubramanian, and P. Sterling, “How much the eye tells the brain”. *Current biology: CB*, vol. 16, no. 14, pp. 1428–34, Jul., 2006.
- [14] A. L. Rothenstein and J. K. Tsotsos, “Attention links sensing to recognition”. *Image and Vision Computing*, vol. 26, no. 1, pp. 114–126, Jan. 2008.
- [15] A. Borji and L. Itti, “State-of-the-art in visual attention modeling”. *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 1, pp. 185–207, Jan. 2013.
- [16] W. Yin, J. Luo, and C. W. Chen, “Event-Based Semantic Image Adaptation for User-Centric Mobile Display Devices”. *IEEE Transactions on Multimedia*, vol. 13, no. 3, pp. 432–442, Jun. 2011.
- [17] J. Baltazar and P. Pinho, “Visual Attention Driven Image to Video Adaptation”. 2005.
- [18] K. Berkner and B. Erol, “Adaptation of document images to display constraints”.
- [19] L.-Q. Chen, X. Xie, X. Fan, W.-Y. Ma, H.-J. Zhang, and H.-Q. Zhou, “A visual attention model for adapting images on small displays”. *Multimedia Systems*, vol. 9, no. 4, pp. 353–364, Oct. 2003.
- [20] M. Hayhoe and D. Ballard, “Eye Movements in Natural Behavior”. vol. 9, pp. 188–194, 2005.
- [21] A. Treisman, “Features and objects in visual processing”. in *Scientific American*, 1986, vol. 255, pp. 114–125.

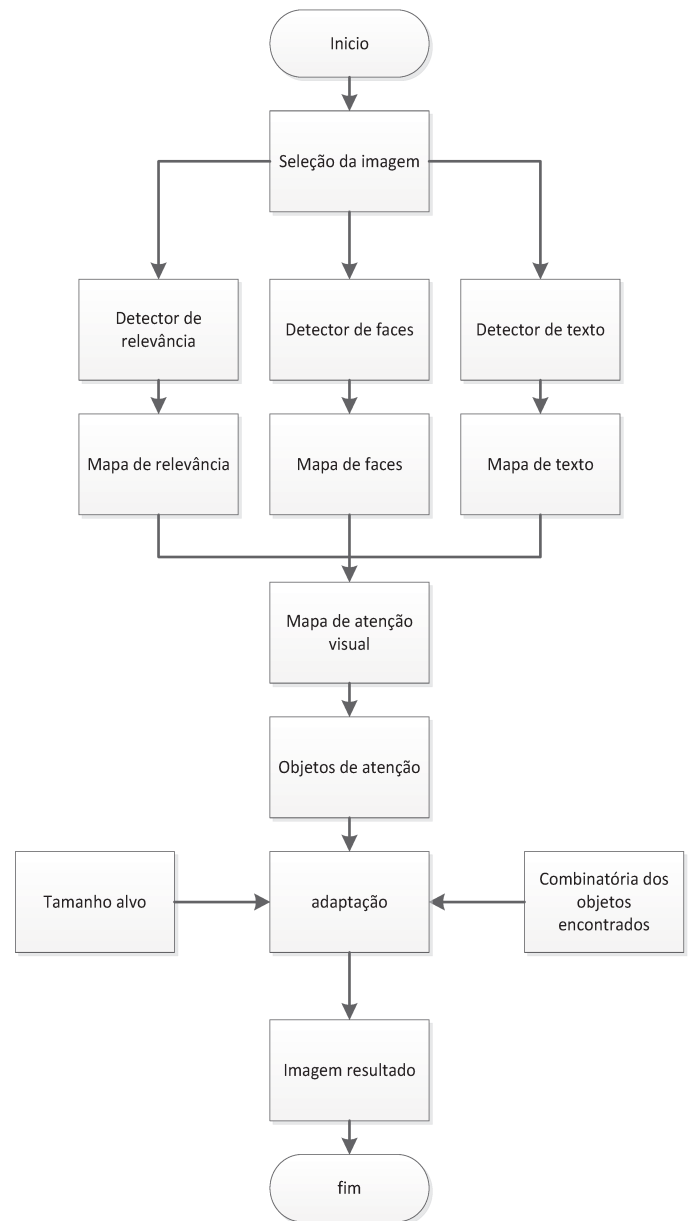
- [22] C. Koch and S. Ullman, “*Shifts in selective visual attention: Towards the Underlying Neural Circuitry*”. *Human Neurobiology* vol 4, no 4. pp. 219–227, 1985.
- [23] C. Christopoulos, A. Skodras, T. Ebrahimi, and C. Unit, “*The Jpeg2000 Still Image Coding System : an Overview*,” vol. 46, no. 4, pp. 1103–1127, 2000.
- [24] P. Sinha, “*Identifying perceptually significant features for recognizing faces*”. pp. 12–21, Jun. 2002.
- [25] L. Itti, C. Koch, and E. Niebur, “*A Model of Saliency-based Visual Attention for Rapid Scene Analysis*”. *IEEE Trans Pattern Analysis Mach Intell*, pp. 1254–1259, 1998.
- [26] Y.-F. Ma, L. Lu, H.-J. Zhang, and M. Li, “*A user attention model for video summarization*”. *Proceedings of the tenth ACM international conference on Multimedia - multimedia '02*, p. 533, 2002.
- [27] P. Viola and M. Jones, “*Robust Real-time Object Detection*”. 2001.
- [28] B. Epshtein, E. Ofek, and Y. Wexler, “*Detecting text in natural scenes with stroke width transform*”. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, no. d, pp. 2963–2970, Jun. 2010.
- [29] X. Tang, W. Luo, and X. Wang, “*Content-Based Photo Quality Assessment*”. *IEEE Transactions on Multimedia*, pp. 1–1, 2013.
- [30] alZahir, S, “*A perceptually perfect image resizing scheme*”. *Digest of Technical Papers. International Conference on Consumer Electronics*, 2005.
- [31] L. Zhang, M.H. Tong, and G.W. Cottrell, “*Sunday: Saliency, Using Natural Statistics for Dynamic Analysis of Scene*”. *Proc. 31st Ann. Cognitive Science Soc. Conf.*, 2009.
- [32] E. Gu, J. Wang, and N.I. Badler, “*Generating Sequence of Eye Fixations Using Decision-Theoretic Attention Model*”. *Proc. Workshop Attention and Performance in Computational Vision*, pp. 277-29, 2007.

- [33] Y. Li, Y. Zhou, J. Yan, and J. Yang, “*Visual Saliency Based on Conditional Entropy*”. *Proc. Ninth Asian Conf. Computer Vision*, 2009.
- [34] J. Yan, J. Liu, Y. Li, and Y. Liu, “*Visual Saliency via Sparsity Rank Decomposition*”. *Proc. IEEE 17th Int’l Conf. Image Processing*, 2010.
- [35] J. Harel, C. Koch, and P. Perona, “*Graph-Based Visual Saliency*”. *Neural Information Processing Systems*, vol. 19, pp. 545-552, 2006.
- [36] W. Kienzle, M.O. Franz, B. Schölkopf, and F.A. Wichmann, “*Center-Surround Patterns Emerge as Optimal Predictors for Human Saccade Targets*”. *J. Vision*, vol. 9, pp. 1-15, 2009.F
- [37] Sung e Poggio, “*Example-Based learning for view-based human face detection*”. *A.I. Memo 1521, MIT Artificial Intelligence Laboratory*, 1994.
- [38] Rowley, Baluha and T. Kanade. “*Neural Network-based face detection*”. In *IEEE T. Pattern Analysis and Machine Intelligence*, 1998.
- [39] Heisle, B., T. Serre, S. Mukherjee and T. Poggio. (2001) “*Feature Reduction and Hierarchy of Classifiers for Fast Object Detection in Video Images*”. In: *Proceedings of 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)*, IEEE Computer Society Press, Jauai, Hawaii, December 8–14, 2001.
- [40] E. Schwartz. “*Spatial mapping in primate sensory projection:analytic structure and relevance to perception*”. *Biological Cybernetics*, 25:181C194, 1977.
- [41] Ara V. Nefian. “*Statistical approaches to face recognition*”, 1996.
- [42] Carles M. “*Nuevos metodos de Analise Multivariante*”,CMC Editions, 2014
- [43] A.G. Leventhal, “*The Neural Basis of Visual Function*”. *Visual and Visual Dysfunction Vol.4*, Boca Raton, 1991.
- [44] Tie Liu, Jian Sun, Nan-Ning Zheng, Xiaoou Tang, Heung-Yeung Shum. “*Learning to Detect A Salient Object*”. *IEEE CVPR 2007*.

Apêndice 1 - Fluxogramas dos programas



a. Modelo Base [19].



b. Modelo Proposto.

Apêndice 2 - Códigos com detalhes

```
function r = VACenter(imgName)

%Detector de relevância

img = imread(imgName);
%To compute a GBVS map:
map = gbvs(img); % map.master_map contains the actual saliency map
%To compute an Itti-Koch-Niebur map:
%map = ittikochmap(img); % map_itti.master_map contains the actual saliency map

[m, n, z]=size(img);

%Detector de faces

map_face=zeros(m,n);
I=img;
faceDetector = vision.CascadeObjectDetector; % Default: finds faces
bbox = step(faceDetector, I); % Detect faces

if bbox>0
    map_face(bbox(1,2):bbox(1,2)+bbox(1,4),bbox(1,1):bbox(1,1)+bbox(1,3))=1;
end
% Draw bounding boxes
shapeInserter = vision.ShapeInserter(...
    'BorderColor','Custom',...
    'CustomBorderColor',[0 255 0]);
I_faces = step(shapeInserter, I, int32(bbox));

%figure, show_imgnmap( I_faces , map_face );
%figure, imshow(I_faces);
map_saliencia=map.master_map_resized;

umb=graythresh(map_saliencia);
bw=im2bw(map_saliencia,0.5);

[L Ne]=bwlabel(bw);
%% Calcular propiedades de los objetos de la imagen
propied= regionprops(L);

imshow(img);
hold on
for n=1:size(propied,1)
    %rectangle('Position',propied(n).BoundingBox,'EdgeColor','r','LineWidth',2)
end

%Detector de texto

% Read image
imagen = img;
% Show image
figure, imshow(img);
% Convert to gray scale
if size(imagen,3)==3 %RGB image
```

```

    imagen=rgb2gray(imagen);
end
% Convert to BW
threshold = graythresh(imagen);
imagen = ~im2bw(imagen,threshold);
% Remove all object containing fewer than 30 pixels
imagen = bwareaopen(imagen,30);
%Storage matrix word from image
word=[ ];
re=imagen;
[m, n z]=size(re);
map_text=zeros(m,n);

%Opens text.txt as file for write
fid = fopen('text.txt', 'wt');
% Load templates
load templates
global templates
% Compute the number of letters in template file
num_letras=size(templates,2);
while 1
    %Fcn 'lines' separate lines in text
    [fl re]=lines(re);
    imgn=fl;
    %imshow(imgn);
    %Uncomment line below to see lines one by one
    %imshow(fl);pause(0.5)
    %-----
    % Label and count connected components
    [L Ne] = bwlabel(imagen);
    for n=1:Ne
        [r,c] = find(L==n);
        % Extract letter
        n1=imagen(min(r):max(r),min(c):max(c));
        [a b]=size(n1);
        [ai bi] = size(imagen);

        if (a*b)<(ai*bi)*3/4
            % Resize letter (same size of template)
            %figure,imshow(n1);
            img_r=imresize(n1,[42 24]);

            %Uncomment line below to see letters one by one
            %imshow(img_r);pause(0.5)
            %-----
            % Call fcn to convert image to text
            letter=read_letter(img_r,num_letras);
            %h = imrect(gca, [min(c) min(r) max(c)-min(c) max(r)-min(r)]);
            rect = [min(c) min(r) max(c)-min(c) max(r)-min(r)];
            if rect(3)>0 & rect(4)>0
                rectangle('Position',[min(c) min(r) max(c)-min(c) max(r)-
min(r)], 'EdgeColor','b', 'LineWidth',2);
            end
            map_text(min(r):max(r),min(c):max(c))=1;
            % Letter concatenation
            word=[word letter];
        end
    end
end

```

Apêndice 3 - Resultados

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
1	282	242	292	270	287	264	287.0	258.7	327.5	154.5	363.5	260	111.76	76.51
2	129	210	122	208	120	141	123.7	186.3	179.5	78	371.5	267	121.87	260.63
3	231	213	257	218	257	217	248.3	216.0	212	101.5	379.5	252.5	120.13	136.15
4	284	223	336	236	189	167	269.7	208.7	200.5	120	43.5	214	112.45	226.23
5	154	161	164	168	161	165	159.7	164.7	182	159	182	159	23.04	23.04
6	174	258	173	263	172	264	173.0	261.7	109.5	48	286	273.5	222.90	113.62
7	107	134	102	125	102	127	103.7	128.7	239	273.5	240	168	198.22	141.89
8	155	152	157	150	152	156	154.7	152.7	200.5	155.5	397	32	45.92	270.71
9	316	191	313	191	296	178	308.3	186.7	260.5	52	395	252.5	142.91	108.84
10	127	96	123	98	120	71	123.3	88.3	172.5	100.5	397.5	82	50.65	274.24
11	221	251	260	276	173	184	218.0	237.0	169	212.5	246	345.5	54.78	112.05
12	201	184	208	192	211	190	206.7	188.7	150.5	240	296	60	76.09	156.64
13	212	208	228	228	226	226	222.0	220.7	200.5	150.5	395	219	73.39	173.01
14	242	171	265	205	260	202	255.7	192.7	255.5	241	187	175.5	48.33	70.78
15	173	214	209	246	188	219	190.0	226.3	167	126	358.5	170.5	102.94	177.51
16	243	204	264	258	264	261	257.0	241.0	285.5	137	283.5	91.5	107.83	151.83
17	245	132	242	283	239	280	242.0	231.7	187	208.5	292	14	59.68	223.34
18	198	209	198	244	192	212	196.0	221.7	196.5	140.5	225	270	81.17	56.37
19	208	171	219	181	217	180	214.7	177.3	197.5	168.5	197.5	168.5	19.31	19.31
20	158	190	127	185	109	189	131.3	188.0	182.5	158.5	259	279	59.06	156.78

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
21	231	271	239	275	175	133	215.0	226.3	133.5	232	148	59.5	81.70	179.78
22	145	256	145	261	143	259	144.3	258.7	169	93.5	166	247	167.00	24.61
23	194	159	198	176	196	176	196.0	170.3	163	220.5	255.5	19.5	60.05	162.14
24	296	205	346	280	346	278	329.3	254.3	264.5	80	275.5	26	186.00	234.59
25	169	252	174	245	173	288	172.0	261.7	65	52	185	262	235.39	13.00
26	216	234	209	240	209	239	211.3	237.7	143	91	143	91	161.80	161.80
27	255	257	241	269	240	267	245.3	264.3	164.5	141	164.5	141	147.46	147.46
28	326	151	343	154	340	154	336.3	153.0	184.5	150.5	365.5	20	151.85	136.16
29	221	98	219	120	209	110	216.3	109.3	108.5	171.5	294	131.5	124.47	80.77
30	168	222	173	235	170	234	170.3	230.3	185.5	165.5	235	238.5	66.58	65.18
31	160	148	157	144	155	165	157.3	152.3	284.5	54.5	284.5	54.5	160.45	160.45
32	130	125	131	124	128	120	129.7	123.0	172.5	66.5	343.5	64	70.90	221.82
33	257	245	125	221	103	155	161.7	207.0	218	189.5	218	189.5	58.99	58.99
34	104	280	105	343	102	355	103.7	326.0	148.5	211	148.5	211	123.43	123.43
35	130	178	140	212	198	259	156.0	216.3	171	164.5	396.5	29	53.96	304.85
36	170	257	196	262	177	219	181.0	246.0	274.5	233.5	274.5	233.5	94.33	94.33
37	130	211	133	194	232	291	165.0	232.0	126	118.5	395.5	188	120.01	234.66
38	191	230	223	244	223	242	212.3	238.7	329	286.5	329.5	94	126.09	186.16
39	240	145	310	142	311	119	287.0	135.3	266.5	217.5	268	13.5	84.69	123.31
40	156	125	156	130	155	125	155.7	126.7	231	59.5	286.5	125	100.93	130.84
41	211	226	225	239	223	239	219.7	234.7	200.5	181.5	376	38.5	56.52	250.84
42	116	178	171	167	111	166	132.7	170.3	156	133	293	14	44.03	223.94

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
43	253	257	262	265	260	263	258.3	261.7	252	195.5	252	195.5	66.47	66.47
44	130	196	134	213	130	183	131.3	197.3	243	155	398	152.5	119.42	270.41
45	265	212	115	193	115	194	165.0	199.7	184	159	382.5	76.5	44.89	249.95
46	201	116	219	139	215	109	211.7	121.3	203.5	142.5	203.5	142.5	22.69	22.69
47	185	262	171	233	169	231	175.0	242.0	275	192.5	395.5	262.5	111.58	221.45
48	168	84	222	103	195	77	195.0	88.0	150.5	95.5	150.5	330	45.13	246.06
49	204	289	211	304	210	273	208.3	288.7	140.5	131.5	263	119.5	171.18	177.78
50	137	207	143	216	138	195	139.3	206.0	158	108.5	246	218	99.27	107.34
51	212	249	222	261	220	254	218.0	254.7	91.5	217.5	275.5	117	131.85	149.19
52	187	246	200	249	199	249	195.3	248.0	263	76.5	288	40	184.37	227.71
53	173	181	174	184	157	171	168.0	178.7	200.5	150.5	26	172	43.01	142.16
54	124	117	173	136	175	124	157.3	125.7	203	150.5	200.5	120	51.98	43.54
55	115	108	177	137	148	98	146.7	114.3	208.5	180	208.5	180	90.20	90.20
56	129	117	147	156	121	111	132.3	128.0	160.5	158.5	223.5	217.5	41.52	127.76
57	200	327	213	332	211	326	208.0	328.3	150.5	214	150.5	214	127.98	127.98
58	225	162	228	165	227	163	226.7	163.3	150.5	270.5	290	288	131.48	139.83
59	225	270	237	293	236	301	232.7	288.0	150.5	217	150.5	217	108.59	108.59
60	187	133	176	137	172	128	178.3	132.7	200.5	187	200.5	187	58.68	58.68
61	259	198	274	182	270	182	267.7	187.3	150.5	213.5	292	10	120.05	179.00
62	222	98	224	94	223	92	223.0	94.7	69.5	277	394	216	238.34	209.67
63	295	209	344	292	343	291	327.3	264.0	158.5	183.5	158.5	183.5	187.04	187.04
64	229	273	191	249	159	209	193.0	243.7	126	359.5	286.5	372	133.81	158.78

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
65	134	136	136	252	134	137	134.7	175.0	282	116	390.5	114	158.71	263.01
66	113	250	135	333	107	239	118.3	274.0	165.5	126	178.5	326.5	155.33	79.85
67	242	187	288	247	283	243	271.0	225.7	259.5	149	321.5	206.5	77.52	54.01
68	274	222	281	233	195	112	250.0	189.0	293	218	350.5	205.5	51.87	101.85
69	236	237	243	242	241	241	240.0	240.0	123.5	177.5	123.5	177.5	132.21	132.21
70	291	124	331	130	332	129	318.0	127.7	200.5	134	200.5	134	117.67	117.67
71	212	277	286	292	287	292	261.7	287.0	105	295	279	346	156.87	61.49
72	195	234	197	278	192	249	194.7	253.7	186	150.5	186	150.5	103.53	103.53
73	290	149	290	146	289	145	289.7	146.7	200.5	211.5	352	107.5	110.25	73.62
74	164	384	165	376	164	383	164.3	381.0	150.5	293.5	167.5	142.5	88.59	238.52
75	254	224	251	242	248	241	251.0	235.7	199	96.5	315	178	148.56	86.15
76	228	270	277	291	275	290	260.0	283.7	176	117	176	117	186.64	186.64
77	166	218	168	223	165	218	166.3	219.7	134	244.5	213	355	40.77	143.15
78	126	193	137	214	133	214	132.0	207.0	186	237.5	266	234.5	62.02	136.79
79	318	170	334	171	332	170	328.0	170.3	90	39.5	237	150	271.59	93.24
80	143	187	153	190	153	188	149.7	188.3	194	196.5	194	196.5	45.08	45.08
81	233	188	248	188	225	185	235.3	187.0	243	132	302.5	293	55.53	125.49
82	203	293	226	301	222	297	217.0	297.0	179.5	209	179.5	209	95.66	95.66
83	69	189	71	192	67	169	69.0	183.3	145.5	128.5	145	236	94.12	92.47
84	259	181	258	182	256	180	257.7	181.0	246.5	161	246.5	161	22.91	22.91
85	201	221	206	228	203	226	203.3	225.0	180	156	180	156	72.84	72.84
86	172	323	135	304	105	306	137.3	311.0	150.5	216	150.5	216	95.91	95.91

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
87	169	164	168	165	167	162	168.0	163.7	197	146	197	146	33.96	33.96
88	254	223	263	253	261	252	259.3	242.7	235.5	120.5	233.5	74.5	124.47	170.14
89	113	210	133	286	80	240	108.7	245.3	231.5	170	231.5	170	144.09	144.09
90	147	294	150	296	150	314	149.0	301.3	233.5	230	233.5	230	110.58	110.58
91	104	156	104	141	103	140	103.7	145.7	178.5	199	193	346.5	91.89	219.81
92	125	181	116	100	95	71	112.0	117.3	113	236.5	336.5	158	119.17	228.15
93	321	164	352	181	330	169	334.3	171.3	182	134	385	78.5	156.84	105.76
94	250	217	252	213	252	231	251.3	220.3	201	46.5	395	63.5	180.97	212.69
95	258	231	273	248	274	246	268.3	241.7	200.5	164.5	387.5	260	102.74	120.57
96	168	252	172	259	173	257	171.0	256.0	60	128.5	288	45.5	169.05	240.83
97	246	148	246	146	245	143	245.7	145.7	249	170.5	393.5	62	25.06	169.87
98	265	127	267	128	267	124	266.3	126.3	221.5	278	395.5	35	158.15	158.20
99	170	171	181	178	180	172	177.0	173.7	200.5	229	232	154	60.12	58.41
100	194	213	188	178	175	181	185.7	190.7	368.5	13.5	368.5	13.5	254.59	254.59
101	142	157	133	228	131	157	135.3	180.7	390.5	135.5	390.5	135.5	259.13	259.13
102	239	274	294	276	294	287	275.7	279.0	373	44	373	44	254.36	254.36
103	261	140	138	133	275	128	224.7	133.7	339	87	339	254	123.49	165.99
104	291	131	291	133	283	120	288.3	128.0	200.5	216	349.5	52	124.33	97.56
105	282	117	295	129	345	128	307.3	124.7	81	143.5	214	153	227.12	97.54
106	136	158	154	167	240	255	176.7	193.3	200.5	150.5	37	262.5	49.02	155.86
107	138	349	149	366	152	357	146.3	357.3	146	185	294	173	172.33	236.19
108	241	277	262	238	260	244	254.3	253.0	266	62.5	284	8	190.86	246.79

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
109	191	196	209	201	205	199	201.7	198.7	200.5	150.5	200.5	150.5	48.18	48.18
110	263	264	263	262	263	261	263.0	262.3	201.5	157	210.5	124.5	121.97	147.49
111	260	233	288	240	288	238	278.7	237.0	155	147.5	185	289.5	152.66	107.38
112	179	314	251	351	250	352	226.7	339.0	158	209.5	158	209.5	146.58	146.58
113	180	204	186	237	224	283	196.7	241.3	165.5	159	165.5	159	88.03	88.03
114	326	126	344	152	165	79	278.3	119.0	216	179	168.5	83.5	86.52	115.43
115	233	158	234	167	232	165	233.0	163.3	241	148	396.5	257.5	17.29	188.68
116	202	129	204	148	201	140	202.3	139.0	200.5	181.5	200.5	181.5	42.54	42.54
117	180	180	183	186	186	186	183.0	184.0	232.5	230	254	135	67.57	86.27
118	221	201	272	208	270	205	254.3	204.7	311.5	139.5	311.5	139.5	86.69	86.69
119	180	255	180	259	178	259	179.3	257.7	151	225.5	151	225.5	42.87	42.87
120	300	217	317	253	313	252	310.0	240.7	226.5	150.5	384	294	122.89	91.22
121	270	221	213	221	297	226	260.0	222.7	295	48.5	292	195	177.65	42.30
122	237	223	391	213	395	239	341.0	225.0	33	105.5	352	254	330.37	31.02
123	181	277	169	199	246	200	198.7	225.3	69	217.5	395.5	211	129.90	197.35
124	166	228	163	232	163	230	164.0	230.0	153	124.5	153	124.5	106.07	106.07
125	270	143	255	126	252	125	259.0	131.3	269	122.5	269	122.5	13.34	13.34
126	242	183	207	141	360	221	269.7	181.7	220.5	114.5	220.5	114.5	83.24	83.24
127	150	273	145	268	144	265	146.3	268.7	104.5	144.5	86.5	269	131.02	59.83
128	323	167	334	166	332	165	329.7	166.0	200.5	142	396.5	33	131.38	148.85
129	168	172	186	181	185	180	179.7	177.7	219.5	139	396.5	190	55.51	217.18
130	298	156	375	155	377	154	350.0	155.0	120	124	395.5	197	232.08	61.92

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
131	123	362	119	271	129	370	123.7	334.3	150.5	200.5	150.5	200.5	136.50	136.50
132	159	303	144	294	140	303	147.7	300.0	122.5	146.5	162	210	155.55	91.13
133	201	268	208	268	202	265	203.7	267.0	150.5	246	255.5	85	57.16	189.24
134	176	162	179	166	173	382	176.0	236.7	142.5	210	153	325.5	42.82	91.76
135	187	179	178	177	177	169	180.7	175.0	200.5	197.5	200.5	197.5	29.99	29.99
136	208	127	193	119	371	297	257.3	181.0	224	139.5	224	139.5	53.23	53.23
137	239	267	238	291	232	286	236.3	281.3	144.5	227.5	145.5	23	106.45	273.84
138	282	216	274	214	273	211	276.3	213.7	243.5	212	243.5	212	32.88	32.88
139	365	148	380	142	376	140	373.7	143.3	228.5	172	228.5	172	147.97	147.97
140	170	176	169	195	168	183	169.0	184.7	311.5	35	351.5	98.5	206.66	201.82
141	166	230	216	260	219	264	200.3	251.3	200.5	150.5	200.5	150.5	100.83	100.83
142	117	354	130	360	124	376	123.7	363.3	165.5	168	165.5	168	199.76	199.76
143	105	237	115	237	108	238	109.3	237.3	134	255	134	255	30.34	30.34
144	257	280	263	277	161	254	227.0	270.3	223.5	248.5	239.5	177	22.11	94.17
145	184	236	182	240	178	238	181.3	238.0	142.5	206.5	242.5	372.5	50.00	147.76
146	238	298	242	300	236	299	238.7	299.0	234	31	289	210.5	268.04	101.81
147	160	329	155	328	149	321	154.7	326.0	154	182.5	154	182.5	143.50	143.50
148	112	190	107	318	105	188	108.0	232.0	150.5	200.5	150.5	200.5	52.90	52.90
149	153	216	162	205	219	400	178.0	273.7	110	230	109.5	46	80.81	237.75
150	151	232	240	236	236	239	209.0	235.7	253.5	335.5	294.5	274	109.30	93.70
151	295	163	295	161	290	161	293.3	161.7	168	121	260.5	143	131.77	37.77
152	144	86	143	86	121	65	136.0	79.0	200.5	188.5	389.5	21	127.08	260.05

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
153	201	242	194	246	194	253	196.3	247.0	154.5	148	154.5	148	107.48	107.48
154	154	80	148	83	145	80	149.0	81.0	200.5	150.5	200.5	150.5	86.50	86.50
155	143	219	144	230	141	388	142.7	279.0	116	234	151	88.5	52.31	190.68
156	212	180	204	195	201	182	205.7	185.7	219	152.5	384.5	53	35.75	222.67
157	197	207	196	216	194	217	195.7	213.3	200.5	128	164.5	269	85.47	63.80
158	234	158	242	167	331	171	269.0	165.3	165	224	354.5	47	119.41	145.99
159	185	116	186	236	170	148	180.3	166.7	200.5	166	391.5	34	20.18	249.38
160	209	194	204	197	205	195	206.0	195.3	48.5	209	313.5	182	158.09	108.32
161	223	243	257	265	167	171	215.7	226.3	183.5	154.5	284	240.5	78.71	69.79
162	222	240	184	186	192	185	199.3	203.7	221	97.5	269.5	247	108.35	82.47
163	96	134	93	135	90	129	93.0	132.7	250	111	362.5	68.5	158.49	277.03
164	191	320	162	207	160	207	171.0	244.7	156.5	182.5	156.5	182.5	63.84	63.84
165	213	274	194	296	115	113	174.0	227.7	200.5	150.5	200.5	150.5	81.59	81.59
166	142	194	139	201	137	199	139.3	198.0	92.5	201	368.5	181	46.93	229.80
167	157	156	153	169	143	165	151.0	163.3	201	151	201	151	51.50	51.50
168	170	182	261	355	245	353	225.3	296.7	151.5	161	151.5	161	154.46	154.46
169	147	219	143	228	138	228	142.7	225.0	200.5	197	388	57	64.25	297.34
170	206	236	190	233	205	259	200.3	242.7	200.5	149.5	396	293.5	93.17	202.16
171	157	369	210	366	203	399	190.0	378.0	130	174	241.5	9.5	212.64	372.08
172	195	292	221	234	139	159	185.0	228.3	160	212.5	293	379.5	29.59	185.78
173	211	259	216	265	208	200	211.7	241.3	237	147	393.5	293.5	97.68	189.17
174	169	276	176	230	195	233	180.0	246.3	133.5	173.5	155.5	287	86.41	47.48

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
175	213	281	219	230	242	246	224.7	252.3	200.5	153	200.5	153	102.23	102.23
176	282	220	295	231	291	229	289.3	226.7	200.5	177.5	166.5	68.5	101.53	200.26
177	166	304	275	325	276	399	239.0	342.7	167.5	73.5	167.5	73.5	278.50	278.50
178	159	218	149	236	158	236	155.3	230.0	200.5	183	66	142	65.18	125.40
179	169	293	160	299	177	300	168.7	297.3	200.5	150.5	273.5	261.5	150.24	110.79
180	117	201	113	200	116	198	115.3	199.7	246.5	324	265	286	180.73	172.78
181	254	198	260	214	260	211	258.0	207.7	250	21.5	392	121	186.34	159.58
182	272	218	283	223	277	218	277.3	219.7	200.5	150.5	273	24	103.38	195.71
183	145	235	144	237	141	233	143.3	235.0	172	206.5	264.5	396.5	40.42	201.90
184	258	242	253	243	359	241	290.0	242.0	180.5	157	180.5	157	138.62	138.62
185	154	225	159	229	156	227	156.3	227.0	260.5	211	260.5	211	105.39	105.39
186	218	126	220	123	217	125	218.3	124.7	241.5	102	241.5	102	32.41	32.41
187	143	210	151	217	150	216	148.0	214.3	200.5	150.5	200.5	150.5	82.65	82.65
188	257	206	247	199	281	233	261.7	212.7	112.5	138.5	295.5	118	166.59	100.53
189	256	124	278	126	272	122	268.7	124.0	200.5	192	372.5	17.5	96.28	148.74
190	206	187	221	195	218	209	215.0	197.0	200.5	134	200.5	134	64.65	64.65
191	141	174	181	184	180	185	167.3	181.0	193	111	387.5	264	74.56	235.29
192	252	226	201	192	197	182	216.7	200.0	200.5	226.5	385	193.5	31.04	168.46
193	165	252	168	229	162	225	165.0	235.3	285	150.5	285	150.5	146.96	146.96
194	241	226	244	225	238	231	241.0	227.3	309	256	386.5	154	73.80	162.94
195	178	235	188	238	188	290	184.7	254.3	191	165.5	380.5	286	89.06	198.38
196	261	219	251	218	256	214	256.0	217.0	55.5	114	55.5	114	225.41	225.41

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
197	175	152	207	139	206	129	196.0	140.0	166	150.5	94	342.5	31.78	226.74
198	81	268	77	264	72	262	76.7	264.7	200.5	147.5	200.5	147.5	170.48	170.48
199	118	295	113	294	113	292	114.7	293.7	135.5	241.5	135.5	241.5	56.17	56.17
200	190	298	183	194	169	235	180.7	242.3	155	216.5	257.5	140	36.42	127.97
201	225	184	82	199	66	205	124.3	196.0	281	253.5	296	280	166.89	191.12
202	203	214	211	218	203	347	205.7	259.7	150.5	200.5	150.5	200.5	80.90	80.90
203	161	167	159	175	294	191	204.7	177.7	205.5	150.5	205.5	150.5	27.18	27.18
204	173	247	173	243	169	241	171.7	243.7	163	148.5	8	292.5	95.56	170.80
205	256	235	259	235	256	235	257.0	235.0	207	118.5	207	118.5	126.78	126.78
206	241	167	244	165	242	162	242.3	164.7	150.5	217.5	150.5	217.5	105.95	105.95
207	147	308	217	309	143	305	169.0	307.3	86	263	215	303	94.10	46.20
208	115	191	110	190	110	195	111.7	192.0	257.5	119.5	257.5	119.5	162.86	162.86
209	210	225	238	244	293	231	247.0	233.3	179.5	115.5	198	146.5	135.80	99.70
210	273	240	250	131	302	238	275.0	203.0	63	279	345.5	217.5	225.21	71.98
211	77	214	70	213	68	211	71.7	212.7	142.5	150.5	370	260.5	94.24	302.14
212	152	280	150	348	148	348	150.0	325.3	204.5	127.5	204.5	127.5	205.20	205.20
213	208	141	207	139	207	130	207.3	136.7	288	156	352	284	82.95	206.48
214	149	366	146	394	280	400	191.7	386.7	56.5	179.5	272.5	326.5	247.36	100.77
215	118	187	93	148	92	145	101.0	160.0	200.5	180	200.5	180	101.49	101.49
216	219	160	218	172	217	172	218.0	168.0	213	182.5	391	211.5	15.34	178.39
217	213	136	286	137	287	135	262.0	136.0	141	92.5	396	144.5	128.58	134.27
218	120	185	123	181	118	179	120.3	181.7	169	268	295	76	99.11	204.14

	Centróides do banco de dados						Centróides comparados						Distância Euclidiana	
	Usuário 1		Usuário 2		Usuário 3		Pro-médio BD		Proposta MP		Modelo Base MB		BD-MP	BD-MB
	x1	y1	x2	y2	x3	y3	x4	y4	x5	y5	x6	y6		
219	168	323	171	370	168	370	169.0	354.3	151	218	278	390.5	137.52	114.84
220	178	211	252	260	251	254	227.0	241.7	227	195	227	195	46.67	46.67
221	151	143	139	141	133	138	141.0	140.7	206.5	105.5	287.5	143.5	74.34	146.53
222	193	167	138	115	270	260	200.3	180.7	242.5	127.5	281.5	245	67.86	103.57
223	152	132	144	123	150	120	148.7	125.0	63.5	308	262.5	329.5	201.85	234.05
224	128	143	183	206	180	203	163.7	184.0	160.5	225	160.5	225	41.12	41.12
225	245	253	247	248	245	247	245.7	249.3	188.5	151	271	89.5	113.74	161.83
226	147	127	124	125	124	123	131.7	125.0	135	56.5	393.5	193	68.58	270.52
227	207	278	157	102	207	276	190.3	218.7	231.5	151.5	269	80.5	78.78	158.99
228	216	285	209	279	232	360	219.0	308.0	150	222	150	222	110.26	110.26
229	222	206	211	204	207	261	213.3	223.7	201.5	97	369	221.5	127.22	155.68
230	242	203	205	199	202	195	216.3	199.0	224.5	177.5	372	20.5	23.00	236.84
231	245	230	249	250	248	248	247.3	242.7	165	130.5	355.5	99	139.14	179.83
232	223	222	228	223	225	221	225.3	222.0	157.5	198	173.5	291	71.95	86.30
233	263	251	149	114	255	248	222.3	204.3	328	157	328	157	115.78	115.78
234	224	214	213	211	207	205	214.7	210.0	196.5	172.5	385	138	41.67	184.93
235	177	190	212	198	198	183	195.7	190.3	158.5	134	359.5	134	67.49	173.25

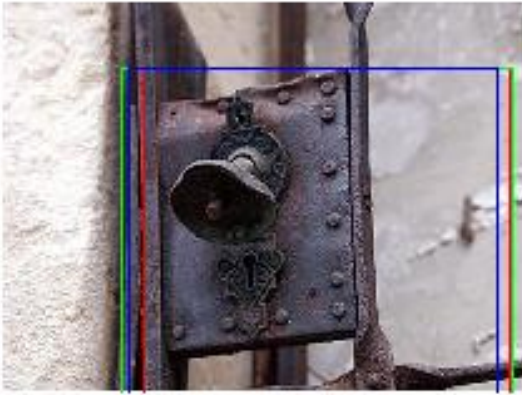
A distância média BD-MP é 107.92

A distância media BD-MB é 140.77

A distância de Mahalanobi BD-MP é 0.71

A distância de Mahalanobi BD-MB é 0.89

Apêndice 4 - Resultados visuais



Banco de Dados

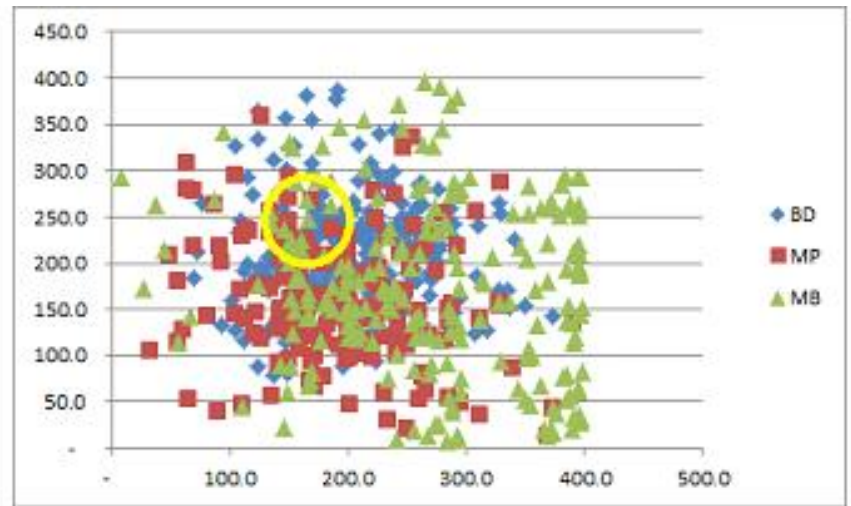


Gráfico de dispersão



Modelo Base



Modelo Proposto



Banco de Dados



Gráfico de dispersão



Modelo Base



Modelo Proposto



Banco de Dados

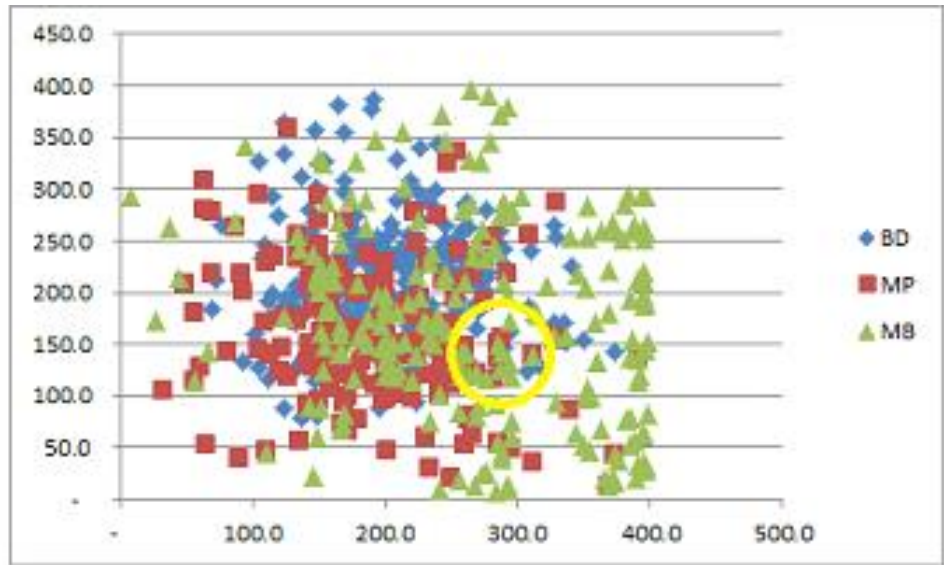


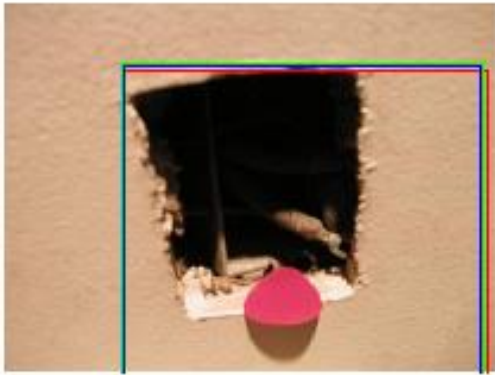
Gráfico de dispersão



Modelo Base



Modelo Proposto



Banco de Dados



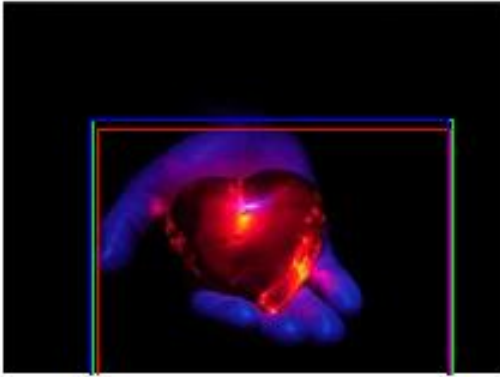
Gráfico de dispersão



Modelo Base



Modelo Proposto



Banco de Dados

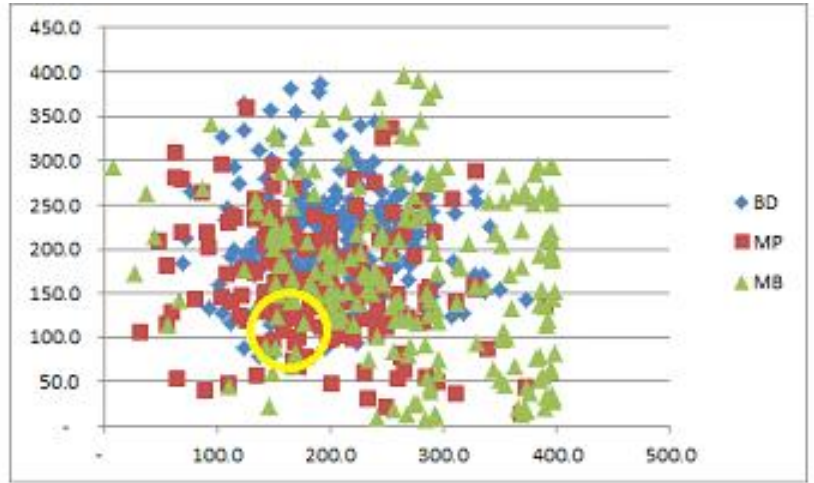
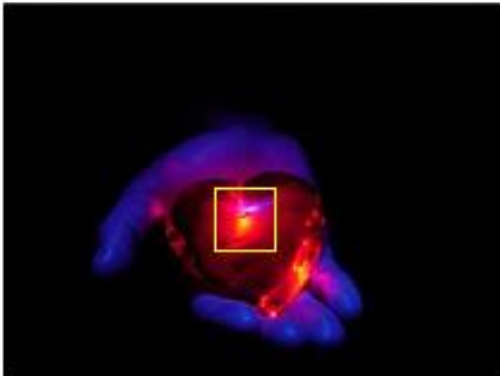
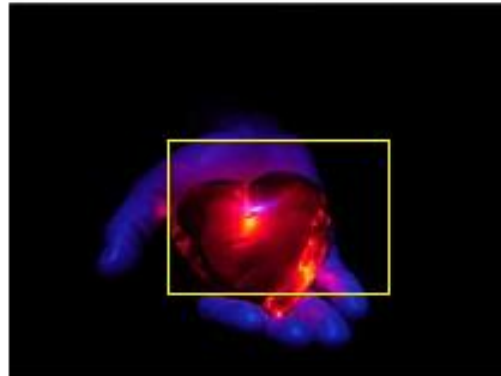


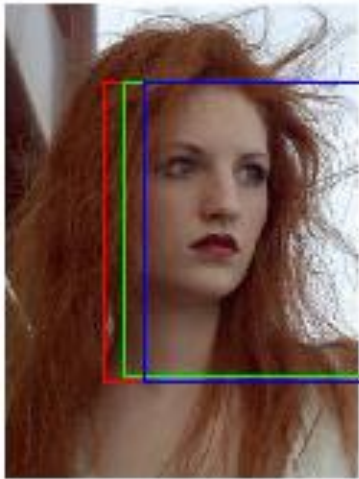
Gráfico de dispersão



Modelo Base



Modelo Proposto



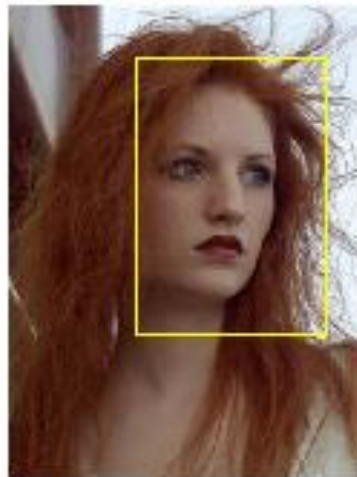
Banco de Dados



Gráfico de dispersão



Modelo Base



Modelo Proposto



Banco de Dados



Gráfico de dispersão



Modelo Base



Modelo Proposto



Banco de Dados



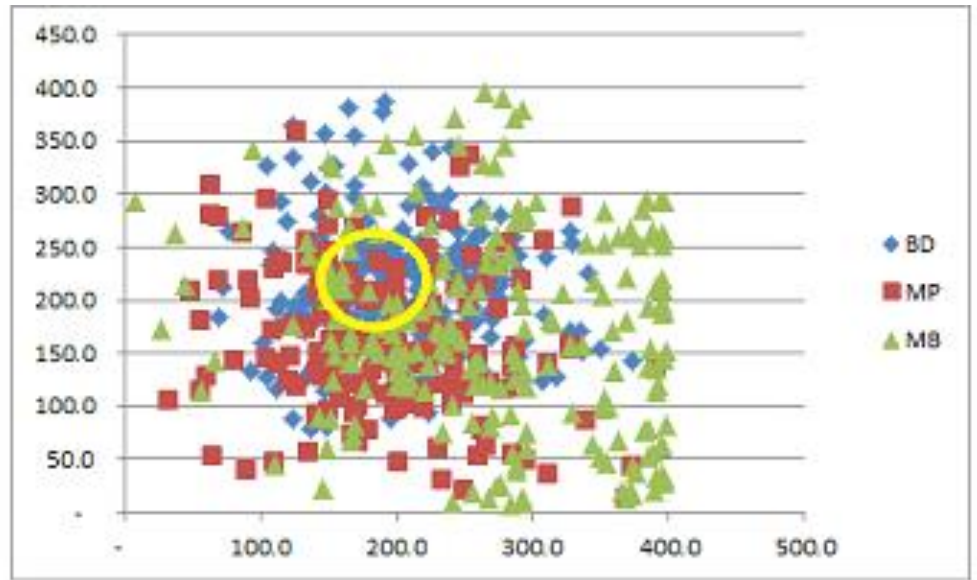
Gráfico de dispersão



Modelo Base



Modelo Proposto



Banco de Dados

Gráfico de dispersão



Modelo Base

Modelo Proposto



Banco de Dados



Gráfico de dispersão



Modelo Base



Modelo Proposto



Banco de Dados



Gráfico de dispersão



Modelo Base



Modelo Proposto