

A Fixation Prediction-based Approach for Image Retargeting Quality Assessment

Saulo Anderson Freitas de Oliveira

Orientador: Prof. Dr. João Paulo Pordeus Gomes

Novembro 2016

1 Introdução

2 Mapas de saliência

3 Bidirectional Fixation Map Similarity

4 Simulações computacionais

5 Conclusão e Trabalhos Futuros

Introdução

O redimensionamento sensível ao conteúdo (**Content-aware Image Retargeting**) é uma classe de técnicas em que se realiza o redimensionamento e, simultaneamente, preservam-se as regiões de interesse nas imagens.



Figura 1: Exemplos de operadores de redimensionamento aplicados em (a) e os respectivos resultados por recorte (CR) em (b), Seam carving (SC) em (c), e Shift-maps (SM) em (d).

- Os algoritmos populares de qualidade de imagem (IQA) tais como, Relação Sinal-ruído de Pico (PSNR), o Índice de Semelhança Estrutural (SSIM), o Índice de fidelidade de informação virtual (VIF), e até mesmo o erro absoluto médio (MAE) não podem ser aplicados diretamente em resultados de redimensionamento porque exigem que os tamanhos das imagens de entrada (original) e de saída (redimensionada) sejam os mesmos.
- Projetar uma métrica de qualidade para retargeting que compara o conteúdo da imagem sob diferentes proporções é significativamente mais desafiador, uma vez que o problema também exige análise semântica e correspondência entre conteúdos.

Avaliação de qualidade em redimensionamento de imagens

-  Área inexistente
-  Pixel
-  Correspondência direta
-  Correspondência semântica

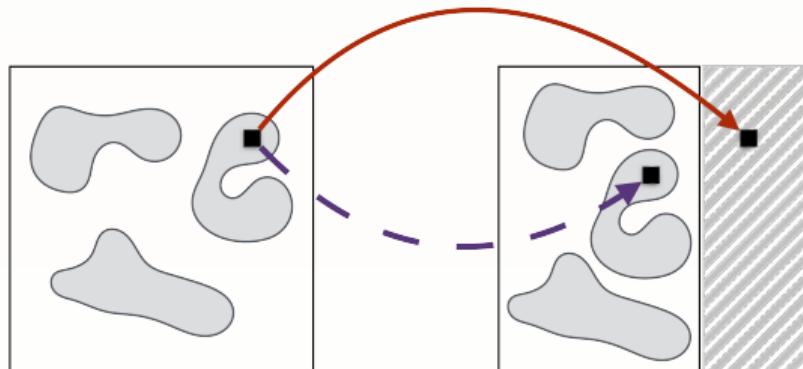
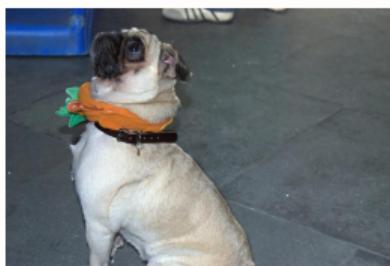


Figura 2: Correspondência de conteúdo após análise semântica em redimensionamento de imagens. A localização espacial dos pixels nem sempre é preservada nos resultados de redimensionamento, por isso a correspondência semântica deve ser empregada para que se possa comparar o conteúdo da imagem.

Mapas de saliência

Mapas de saliência

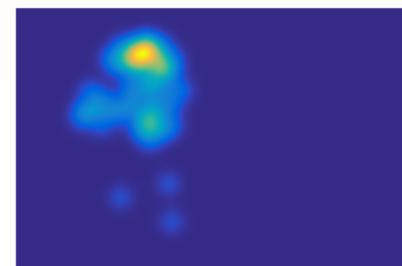
- Mapas de saliência são mapas topográficos das partes visualmente **marcantes** de cenas em que objetos visualmente protuberantes são facilmente destacados.
- Na literatura, um modelo de saliência pode ser usado para duas tarefas diferentes: **detecção de objeto(s) saliente(s)** e **predição de fixação**.



(a) Original.



(b) Detecção de objeto
saliente .



(c) Predição de fixação.

Figura 3: Exemplos de aplicações de modelos de saliências para detecção de objeto saliente em (a) e predição de fixação em (b).

Bidirectional Fixation Map Similarity

- Aplicar predição de fixação para gerar mapas de fixação via SPPM.
- Avaliar a similaridade entre mapas de fixação de modo bidirecional, ou seja, avaliar os pontos de fixação entre ambas direções: da original pra redimensionada e da redimensionada para a original.
- Avaliar correspondência entre pontos de fixação detectados em ambas imagens.
- Combinar os escores de similaridade e a informação de correspondência para gerar o escore final.

Bidirectional Fixation Map Similarity – Visão Geral

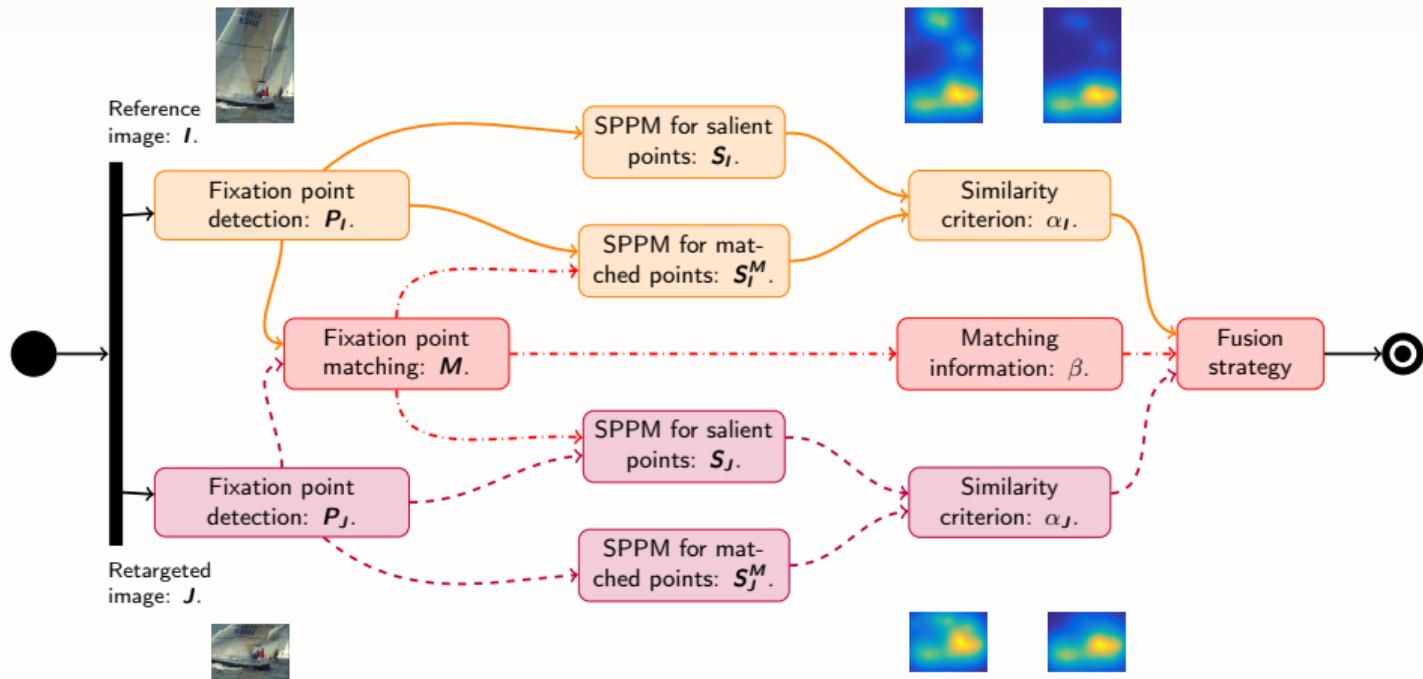
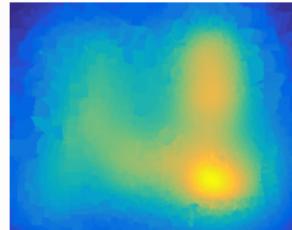


Figura 4: Bidirectional Fixation Map Similarity flowchart.

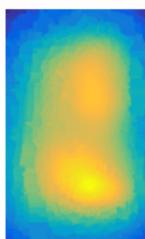
Bidirectional Fixation Map Similarity - Exemplo



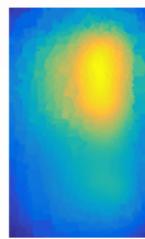
(a) I



(b) S_I

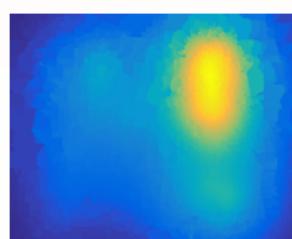


(c) J_{SEAM}



(d) $S_{J_{SEAM}}$

(e) $S_{J_{SEAM}}^M$



(f) S_I^M

Figura 5: Mapas de fixação para imagem Volleyball (a) e o resultado de redimensionamento pelo operador Seam Carving em (c). Escore de qualidade via BFMS ≈ 30 , média da avaliação subjetiva com 34 pessoas, ≈ 29 .

Simulações computacionais

Validação do SPPM como modelo de predição de fixação.

- Dois *datasets*: 1) MIT300 Saliency Benchmark (MIT300) & 2) JuddDB.
- MIT300 não possui rótulos, enquanto que o JuddDB sim.
- Ambos possuem valores de referências (*baselines*) para avaliação de performance, sendo 5 no MIT300 e 2 no JuddDB.
- MIT300 contém 62 modelos, enquanto que no JuddDB existem 41.

Validação do BFMS como métrica para resultados de redimensionamento.

- Um *dataset*: CUHK.
- Avaliação das combinações de critérios de similaridade e estratégias de fusão (combinação) de escores.
- Possui os rótulos e os desempenhos de 11 modelos.

Validação do SPPM como modelo de predição de fixação.

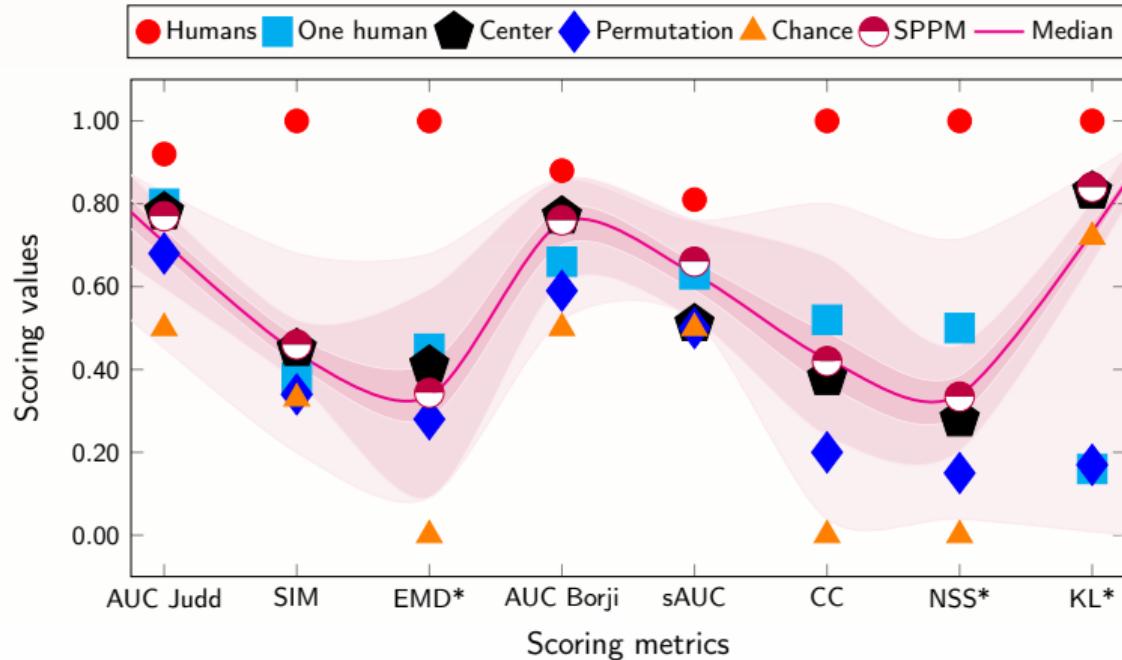


Figura 6: Model performances in MIT300 Saliency Benchmark. The results were taken directly from the benchmark website. We recall that EMD*, NSS* and KL* scores are normalized by range. The brighter area stands for the variation of scoring values for all models in the benchmark, while the darker area is for the 25th and 75th percentiles, and the max and min scoring values.

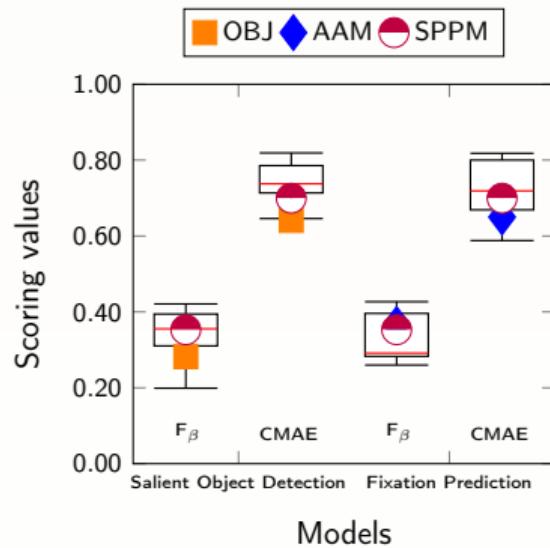


Figura 7: Desempenho dos modelos no JuddDB. Para mostrar o desempenho do SPPM, os resultados foram divididos em duas categorias de modelo: **Salient Object Detection** and **Fixation Prediction**.

- Dois índices de correlação: LCC (linear) e SRCC (ranking/postos).
- Dois índices de erro: RMSE e OR (*outlier ratio*).
- Realizar uma regressão entre os MOS (rótulos) e o valor predito de acordo com

$$f(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{\exp(\beta_2(x - \beta_3))} \right) + \beta_4 x + \beta_5, \quad (1)$$

em que β são parâmetros de regressão que são determinados pela minimização da soma das diferenças quadráticas entre as pontuações (MOS).

Validação BFMS

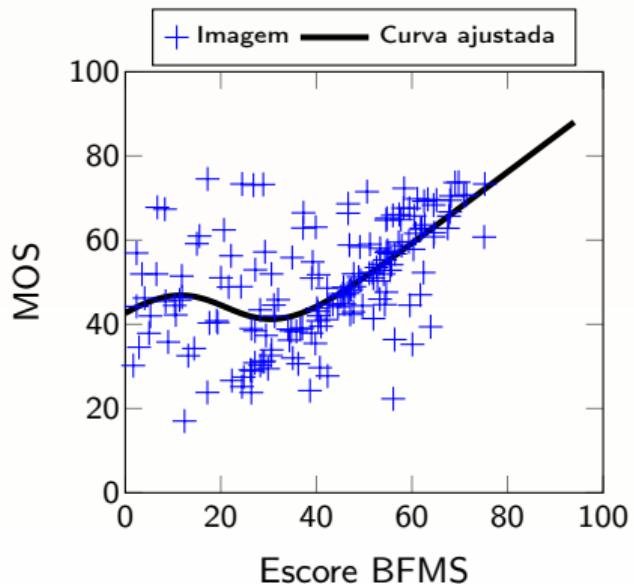


Figura 8: Scatter plot dos MOS *versus* os escores obtidos pelo BFMS.

Resultados

Tabela 1: Combinações de critérios de similaridade e estratégias de fusão em BFMS. ADD significa a adição, MULT a multiplicação, LIN é regressão linear, e NLIN é para a regressão não linear usando *Relevance Vector Machine* (RVM).

Similarity	Fusion	LLC	SRCC	RMSE	OR
CC	ADD	0.4405	0.4900	12.1615	0.0415
	MULT	0.4619	0.4601	11.9881	0.0170
	LIN	0.4461	0.4568	12.0831	0.0292
	NLIN	0.5532	0.5777	11.4908	0.0421
CMAE	ADD	0.4645	0.4734	12.1420	0.0058
	MULT	0.4622	0.4608	11.9764	0.0181
	LIN	0.4413	0.4811	12.1152	0.0404
	NLIN	0.5405	0.5497	11.3423	0.0345
SIM	ADD	0.4982	0.5093	11.8286	0.0170
	MULT	0.4651	0.4601	11.9518	0.0170
	LIN	0.4494	0.5012	12.0690	0.0468
	NLIN	0.5890	0.6222	10.8283	0.0374

Resultados

Tabela 2: Performance comparison on CUHK retargeting database.

IRQA	LLC	SRCC	RMSE	OR
ARS (2016)	0.6835	0.6693	9.8550	0.0702
BFMS	0.5890	0.6222	10.8283	0.0374
PGD (2014)	0.5403	0.5409	11.3610	0.1520
GIST (2015)	0.5443	0.5114	11.3260	0.1579
NR (2016)	0.5371	0.4926	-	0.1928
GLS (2014)	0.4622	0.4760	10.9320	0.1345
CSim (2011)	0.4374	0.4662	12.1410	0.1520
EH (2001)	0.3422	0.3288	12.6860	0.2047
EMD (2009)	0.2760	0.2904	12.9770	0.1696
SIFT-Flow (2011)	0.3141	0.2899	12.8170	0.1462
BDS (2008)	0.2896	0.2887	12.9220	0.2164
PHOW (2007)	0.3706	0.2308	12.5400	0.2222

Conclusão e Trabalhos Futuros

SPPM

- SPPM é competitivo, alcançando a classificação média em dois datasets bem conhecidos. Assim, apoiamos que através da melhoria dos refinamentos automáticos (estratégias internas), a qualidade resultante pode ser melhorada.
- Além disso, em termos de simplicidade, nossa proposta é atrrente porque requer apenas um único parâmetro a ser ajustado, o tamanho da janela, enquanto outras metodologias requerem muitos hiper-parâmetros.

BFMS

- Diferentes combinações de similaridades entre mapas de fixação e estratégias de fusão foram exploradas, a fim de encontrar uma combinação viável em que nossa proposta pudesse tirar melhor proveito.
- Os resultados experimentais indicam que o BFMS é promissor para a avaliação da qualidade. Isso é confirmado pela boa correlação obtida em termos de pontuações subjetivas quando comparadas com os outros algoritmos, nos quais obtivemos o segundo lugar.
- Além disso, os resultados sugerem que a relação linear entre as características extraídas e as pontuações de opinião médias é baixa.

SPPM

- Hoje em dia, busca-se incorporar algumas características *top-down*, tais como a detecção de rosto e de pessoas, a fim de melhorar a abordagem atual.
- Além disso, pretende-se explorar uma maneira automática de escolher valores adequados para o tamanho da janela e os efeitos de usar diferentes funções do kernel.

BFMS

- Incluir outros atributos referentes de modo a melhorar a qualidade do padrão (tornar o conjunto de valores mais representativo), tais como informação estrutural, forma, e escala do redimensionamento, dentre outras.

Obrigado pela atenção!

Saulo A. F. Oliveira

sauloafoliveira@lia.ufc.br

<http://lattes.cnpq.br/9883694006602467>

João Paulo Pordeus Gomes

jpgomes@lia.ufc.br

<http://lattes.cnpq.br/9553770402705512>