Video Quality Assessment in Java

Pós-graduação strictu sensu *em Ciência da Computação*

PROCESSAMENTO DE IMAGENS, Prof. Tsang Ing Ren

Estêvão Monteiro

Ricardo Scholz

Junho de 2014

Centro de Informática UFPE

Universidade Federal de Pernambuco

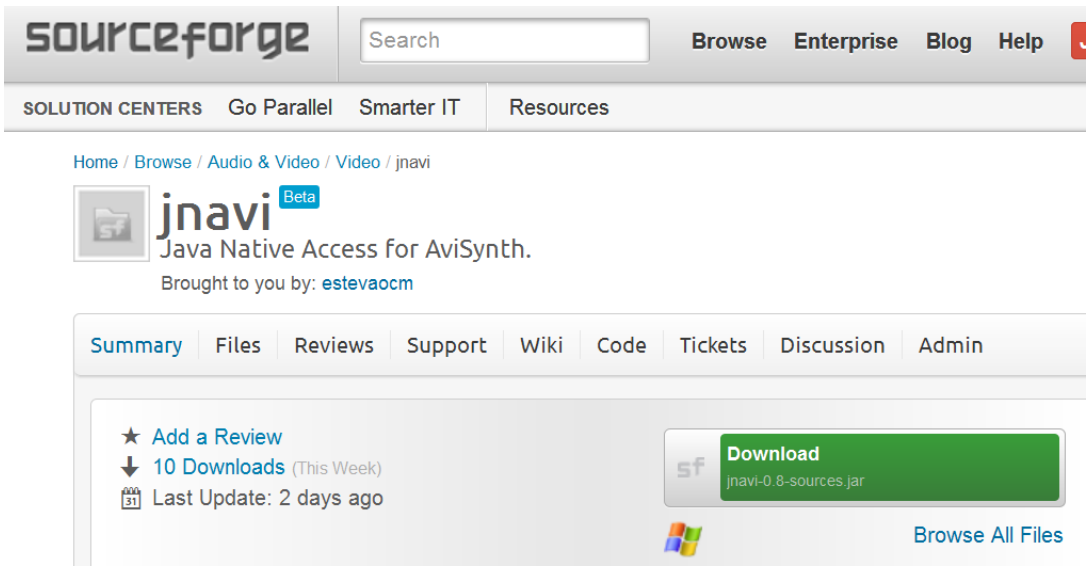
Introdução

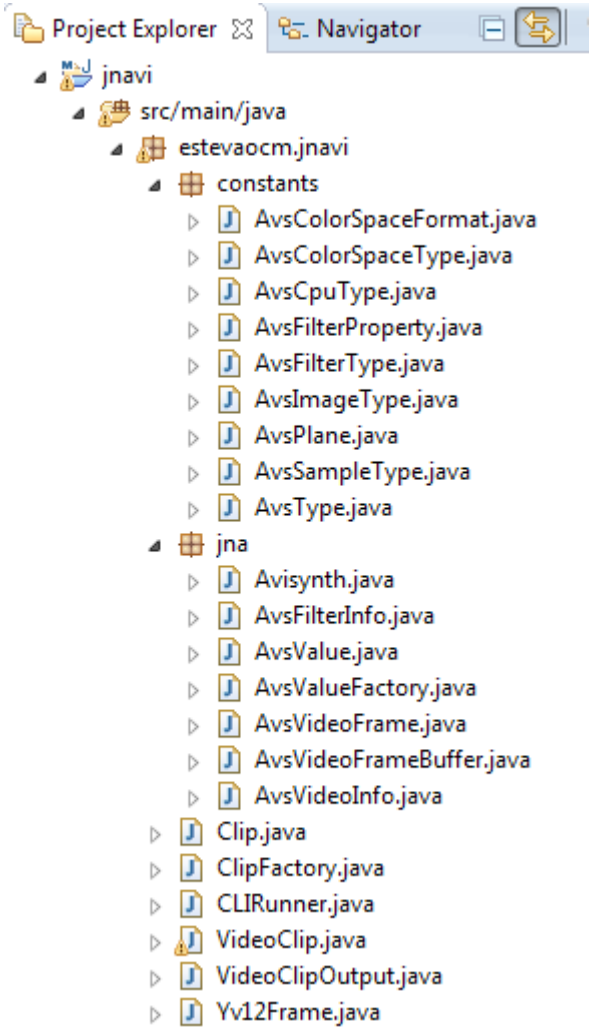
- Fidelidade de vídeo:
  - Bits por pixel por frame
  - Técnicas avançadas de compressão otimizam a alocação de bits
  - Peculiaridades psicovisuais
- Vídeo adaptativo (Youtube, Netflix…):
  - Banda disponível
  - Varia bits por frame (taxa de dados)
  - Varia pixels por frame (resolução)
  - Versões devem balancear taxa de dados x resolução

Avisynth + Java

- Avisynth
  - Entrada: script
  - Saída: frame decodificado pelo SO
  - DLL no Windows
  - Entrada para MSU Video Quality Measurement Tool (benchmark)
  - Entrada para o encoder x264 (H.264, GPL, alta qualidade)
- JNAVI - Java Native Access for Avisynth
  - Encapsula o AviSynth
  - Recupera frames Y'UV12
    - Formato de cores planar, 12 bits
    - Y = luminância, resolução x1 (8 bits)
    - U = diferença entre azul e amarelo, resolução x 0,25 (2 bits)
    - V = diferença entre verde e vermelho, resolução x 0,25 (2 bits)

Avisynth + Java = jnavi





JVQA

- JVQA - Video Quality Assessment for Java
  - Foco em performance (meta: 30 frames/segundo para SD - 300 mil pixels)
  - Implementações originais de operações sobre imagem (sem bibliotecas)
- Métricas
  - Structural Similarity (SSIM): correlação de variância, janela gaussiana [Wang, Bovik et al]
    - Luminância, contraste e estrutura;
  - Gradient SSIM (GSSIM): correlação de gradiente (Sobel), janela gaussiana [Chen, Yang, Xie]
  - Multiscale (MSSSIM): original mais 4 subamostragens pela metade das dimensões [Wang, Bovik, Simoncelli]

JVQA

sourceforge | Search | Browse | Enterprise | Blog | Help | Jobs | Me | 16

SOLUTION CENTERS | Smarter Commerce | Go Parallel | HTML5 | Smarter IT | Resources

Home / Browse / Projects / jvqa / Code

jvqa Alpha

Video Quality Assessment in Java

Brought to you by: estevaocm

Summary | Files | Reviews | Support | Wiki | Code | Tickets | Discussion | Admin

Browse Commits

Fork

Branches

master

Tags

batchtests

true_valid_ssim

valid_ssim

Tree [9a6cef] master / | Download Snapshot | History

RW | RO | HTTP | Read/Write access | git clone ssh://estevaocm@git.code.sf.net/p/jvqa/code jvqa-code

File	Date	Author	Commit
jvqa	1 day ago	estevaocm	[9a6cef] methods normalization

Project Explorer | Navigator

- jvqa [code master]
  - src/main/java
    - estevaocm.jvqa
      - gui
      - image
        - AbstractGradientImage.java
        - CrossCorrelationImage.java
        - CrossCorrelationMap.java
        - GaussFilter.java
        - Image.java
        - ImageConvolver.java
        - ImageDownsampler.java
        - ImageFileSaver.java
        - ImageFileType.java
        - ImageMeansCalculator.java
        - IntegerGaussFilter.java
        - IntegerIntegralImage.java
        - IntegralImage.java
        - MatrixOperations.java
        - RobertsBovikGradientImage.java
        - RobertsBovikGradientIntMatrix.java
        - RobertsGradientImage.java
        - SobelBovikGradientImage.java
        - SobelGradientImage.java
        - TypeConverter.java
      - ssim
        - luminance
          - AbstractLuminanceComparator.java
          - DownscaledLuminanceComparator.java
          - IntegralLuminanceComparator.java
        - structure
          - AbstractStructureComparator.java
          - FullGradientComparator.java
          - FullGradientGaussComparator.java
          - FullVarianceComparator.java
          - FullVarianceGaussComparator.java
          - GaussGradientComparator.java
          - GaussStructureComparator.java
          - GradientGaussComparator.java
          - PlainStructureComparator.java
          - StructureGaussComparator.java
          - VarianceGaussComparator.java
          - VarianceStructureComparator.java
        - AbstractSsimComparator.java
        - AbstractSsimIndex.java
        - DownscaledSsimIndex.java
        - FullscaleSsimIndex.java
        - MultiscaleSsimIndex.java
        - TransformationCache.java

Otimizações

- Inteiros em vez de pontos flutuantes [Chen, Bovik]
- Imagem integral para luminância [Chen, Bovik]
- **Subamostragem da imagem integral [Monteiro, Scholz]**
- Correlação de gradiente de Roberts [Chen, Bovik]
- Janela gaussiana de inteiros [Chen, Bovik]
- **Sem janela, valores crus [Monteiro, Scholz]**
- Tratamento de zeros em vez de constantes estabilizadoras [Rouse, Hemami]
- **Tratamento de zeros diferenciado [Monteiro, Scholz]**

Otimizações: Integral

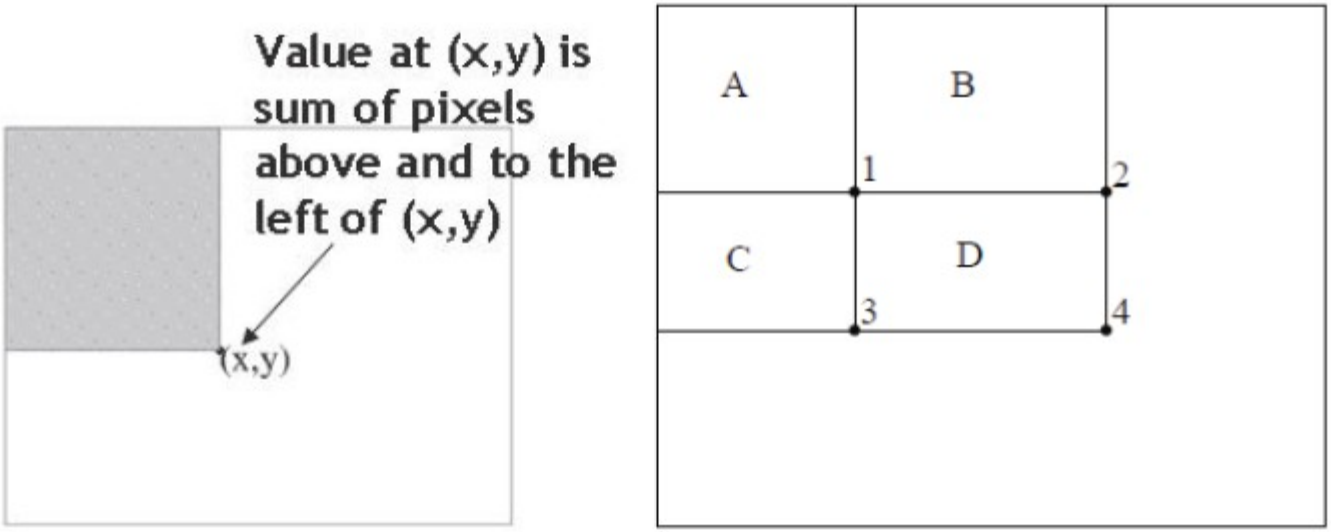
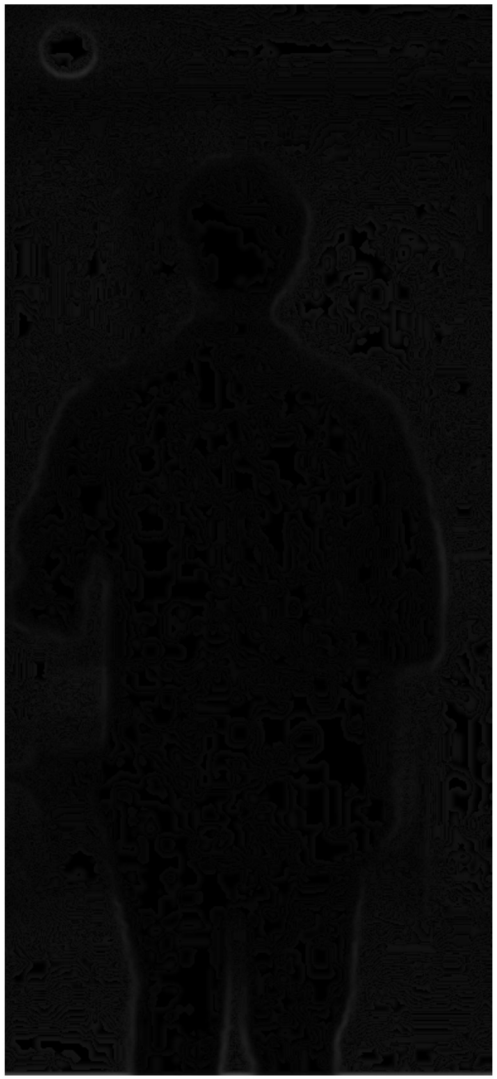


Fig. 1. Left: Integral image. Right: How to compute sum value over region D in integral image domain.

Original x Variância



Otimizações: Gradiente de Sobel

-1	0	+1
-2	0	+2
-1	0	+1

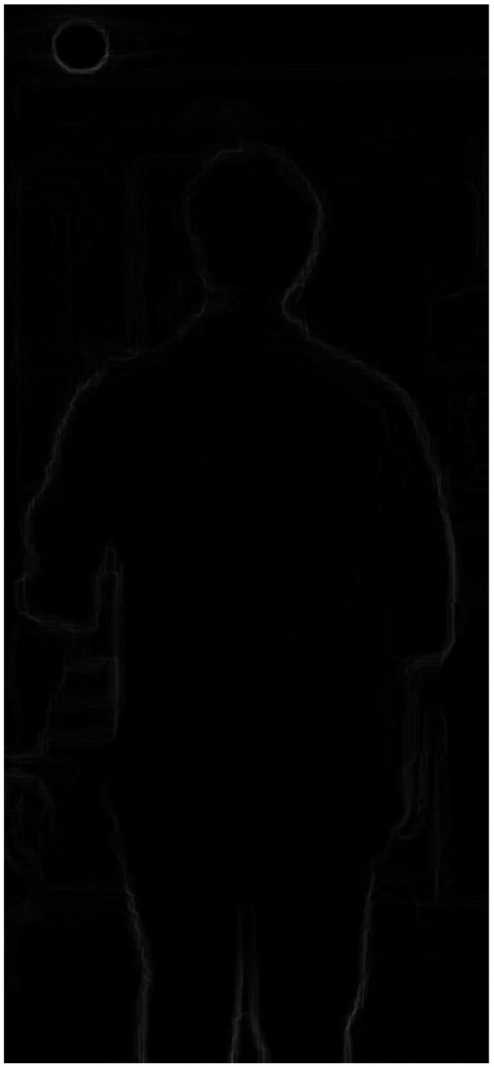
Vertical edge mask

−1	−2	−1
0	0	0
+1	+2	+1

Horizontal edge mask

Fig.2 Sobel operator masks

Otimizações: Gradiente de Roberts



-1	0
0	1

0	-1
1	0

Fig. 2. Roberts gradient templates.

The gradient magnitude is approximated by[1]

$$|\nabla I| = \max\{|\nabla i|, |\nabla j|\} + (1/4)\min\{|\nabla i|, |\nabla j|\}$$

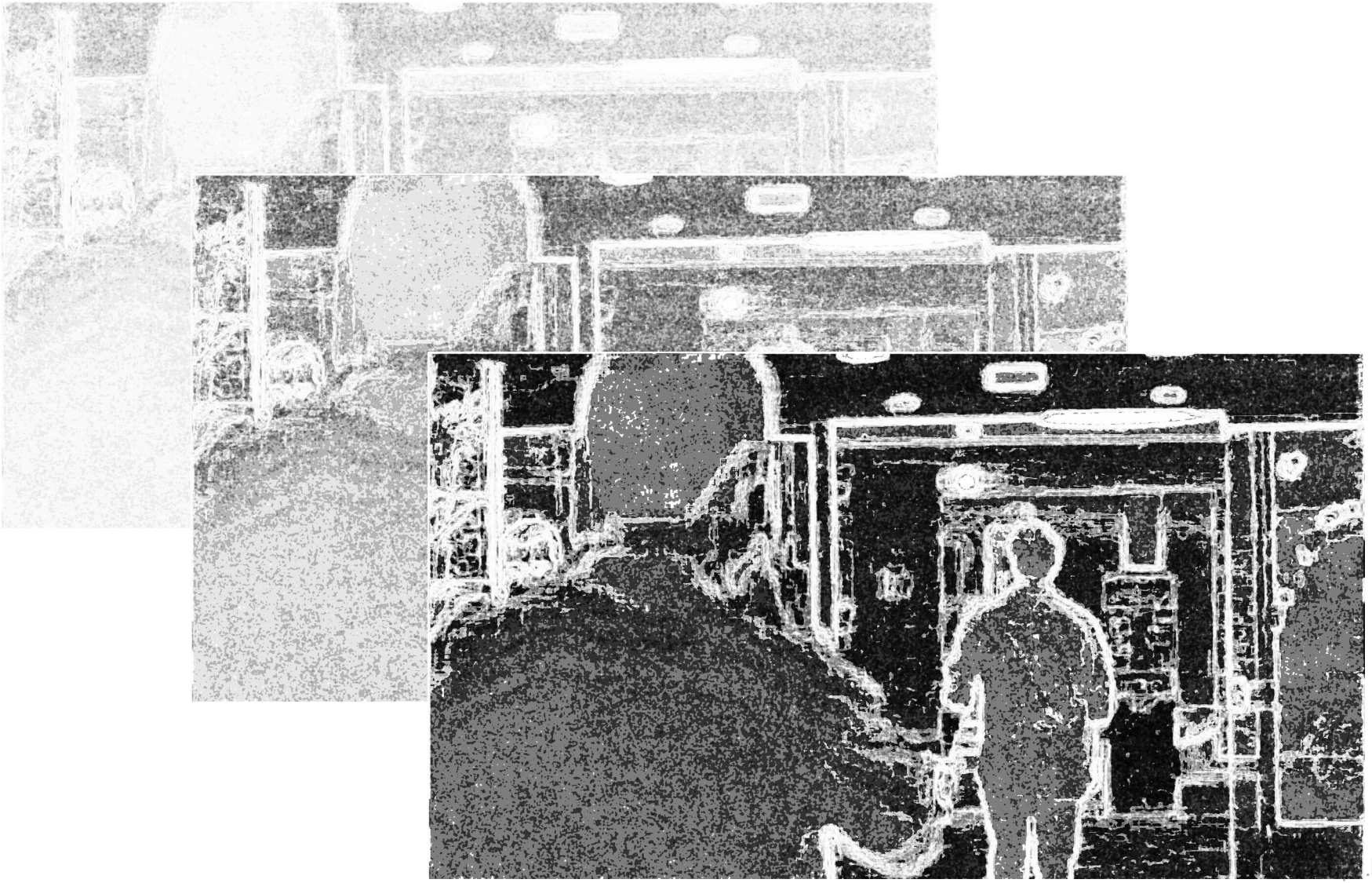
[1] An accurate truncated expansion approximation learned by author ACB whilst lecturing at Texas Instruments in the 1990s.

Otimizações: Janela Gaussiana

0	0	0	1	1	0	0	0
0	0	1	2	2	1	0	0
0	1	2	4	4	2	1	0
1	2	4	8	8	4	2	1
1	2	4	8	8	4	2	1
0	1	2	4	4	2	1	0
0	0	1	2	2	1	0	0
0	0	0	1	1	0	0	0

Fig. 3. 8x8 integer approximation to Gaussian window.

Otimizações: Estabilização



Experimentos

- Vídeo original:
  - 720x404 (subamostragem de Blu-ray)
  - 1min 36s
  - 2300 frames
- Vídeos derivados:
  - x264
  - 600 kbit/s
  - Tunings, em ordem de ranking psicovisual:
    - Psicovisual (quantização adaptativa e retenção de energia visual)
    - SSIM (favorece borramento)
    - PSNR (média de erros quadráticos, favorece borramento)

Experimentos

- 32 combinações:
  - Escala: original x múltiplas
  - Luminância: original x subamostrada
  - Estrutura: variância x gradiente
  - Janela: Gauss 8x8 x nenhum (pixel a pixel)
  - Estabilização de zeros: lógica x constante
  - 3 vídeos x 32 métricas = 96 índices

Experimentos

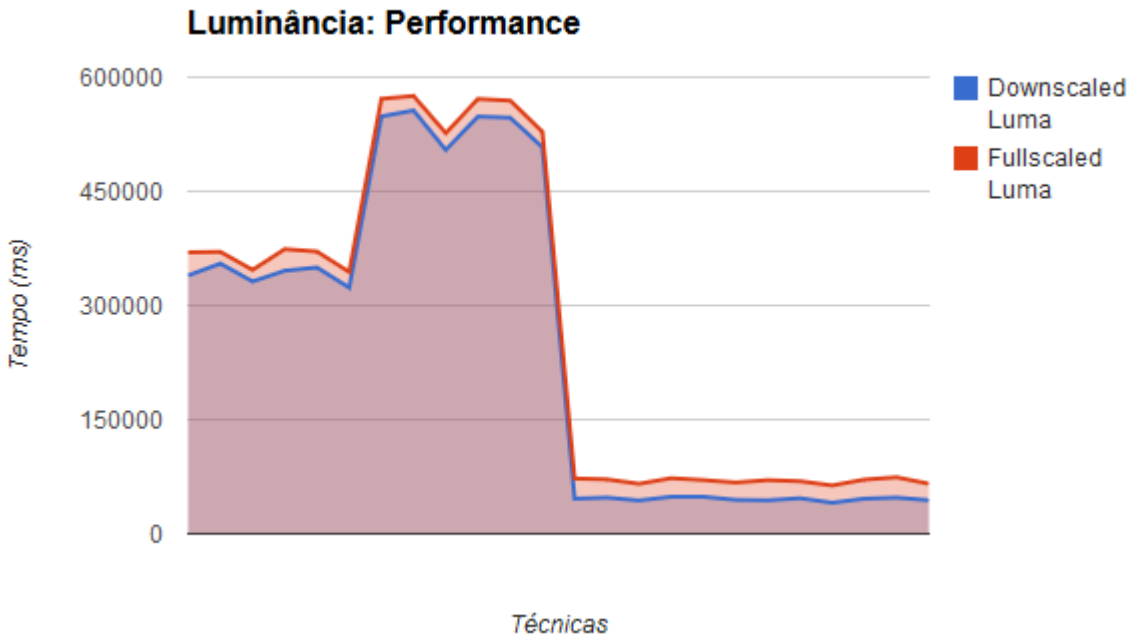
- Dados de interesse
  - Ranking desejado: psicovisual > SSIM > PSNR
  - Performance
  - Amplitude da variação dos índices (maior é melhor para comparações)

- Vazamento de memória no Avisynth
  - Falha após 800 MB lidos: mais de 2000 SD ou mais de 400 Full HD
  - Contornado com subamostragem: a cada 2 frames e a cada 3 frames

Resultados

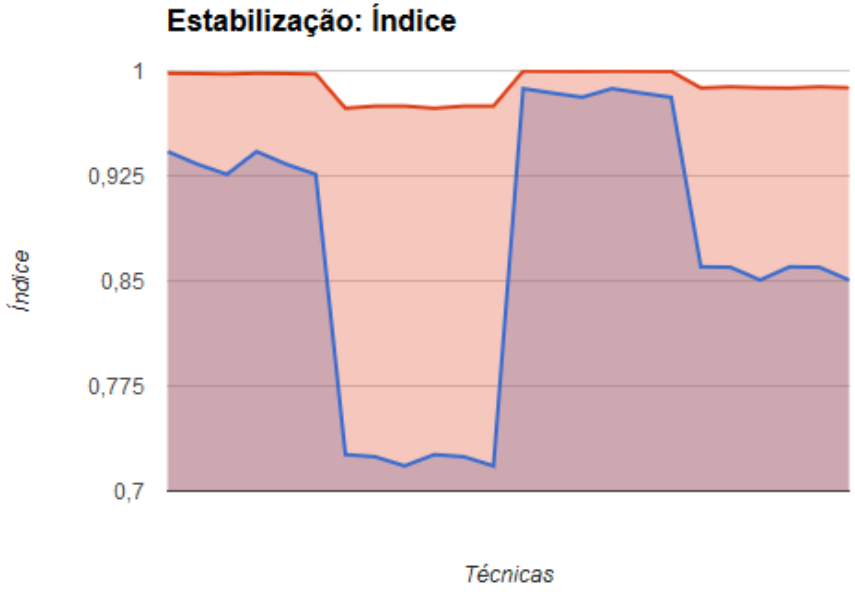
	A	B	C	D	E	F	G	H	I
1	**Fullscale gradient gauss:**	**Psy-grain**	**Time**	**SSIM**	**Time**	**PSNR**	**Time**	**Rank Class**	**Amplitude**
2	fullscale-fastluma-gradient-gauss-zero	**0.8750817696804699**	5'42"390	0.852469028214709	5'44"776	0.8432016446057604	5'47"756	Desired	0,0318801250747101
3	fullscale-fastluma-gradient-gauss-constant	**0.9972018101890915**	5'45"291	0.9968179025176216	5'44"932	0.996468182184468	5'42"452	Desired	0,0007336280046240
4	fullscale-mediumluma-gradient-gauss-zero	**0.8750792519762387**	5'56"570	0.8524664155541609	5'56"944	0.8431989086502581	5'56"196	Desired	0,0318803433259810
5	fullscale-mediumluma-gradient-gauss-constant	**0.9971986322238436**	5'57"225	0.9968145680358038	**5'56"195**	0.9964646280066208	5'57"350	Desired	0,0007340042172220
6	**Fullscale variance gauss:**								
7	fullscale-fastluma-variance-gauss-zero	0.9767376927700855	8'26"284	0.9774660791879395	8'25"472	**0.9774851856728569**	8'25"238	Inverse	0,0007474929027710
8	fullscale-fastluma-variance-gauss-constant	0.9914746069899023	8'26"299	**0.992762005761253**	8'26"751	0.9927601409361219	8'26"595	Expected	0,0012873987713510
9	fullscale-mediumluma-variance-gauss-zero	0.9767349668486592	8'38"420	0.9774632060979411	8'40"12	**0.9774821356929321**	8'41"119	Inverse	0,0007471688442730
10	fullscale-mediumluma-variance-gauss-constant	0.9914714494925885	8'39"512	**0.9927586875486971**	**8'39"75**	0.9927566029052538	**8'38"623**	Expected	0,0012872380561091
11	**Fullscale gradient unfiltered:**								
12	fullscale-fastluma-gradient-unfiltered-zero	**0.6653726307335907**	1'40"277	0.6517500722889028	1'39"856	0.6445999808290491	1'40"12	Desired	0,0207726499045421
13	fullscale-fastluma-gradient-unfiltered-constant	0.9735444125140393	**1'39"583**	0.9755522167148908	**1'39"279**	**0.9756473029708912**	**1'39"450**	Inverse	0,0021028904568520
14	fullscale-mediumluma-gradient-unfiltered-zero	**0.6653707930691323**	1'52"476	0.6517481680717736	1'52"586	0.6445979713060613	1'52"414	Desired	0,0207728217630710
15	fullscale-mediumluma-gradient-unfiltered-constant	0.9735413084034888	**1'52"445**	0.975548950431287	**1'52"429**	**0.9756438198409636**	1'54"145	Inverse	0,0021025114374740
16	**Fullscale variance unfiltered:**								
17	fullscale-fastluma-variance-unfiltered-zero	0.99987098433037	1'40"760	**0.9998796729552794**	1'40"199	0.9998614318195603	1'41"759	Anomaly	0,0000182411357191
18	fullscale-fastluma-variance-unfiltered-constant	0.9998629463138488	1'41"104	**0.9998720907051113**	1'40"682	0.9998526274671645	1'42"539	Anomaly	0,0000194632379470
19	fullscale-mediumluma-variance-unfiltered-zero	0.9998681765436203	1'54"520	**0.9998767147683905**	1'53"444	0.9998582916425578	1'54"520	Anomaly	0,0000184231258319
20	fullscale-mediumluma-variance-unfiltered-constant	0.9998597613389566	1'54"566	**0.9998687482033638**	1'54"520	0.9998490635382301	1'54"535	Anomaly	0,0000196846651341
21	**Multiscale gradient gauss:**								
22	multiscale-fastluma-gradient-gauss-zero	**0.9793113002508889**	7'11"544	0.9749911286019437	7'10"498	0.9714124963206313	7'10"545	Desired	0,0078988039302580
23	multiscale-fastluma-gradient-gauss-constant	**0.999616812558127**	7'11"824	0.9995483419162436	7'11"75	0.999453880311436	7'11"154	Desired	0,0001629322466910
24	multiscale-mediumluma-gradient-gauss-zero	**0.9793102347545912**	**7'10"342**	0.9749898633939332	**7'10"717**	0.9714109885586235	7'11"309	Desired	0,0078992461959671
25	multiscale-mediumluma-gradient-gauss-constant	**0.9996155075317397**	7'11"949	0.9995468035337404	7'11"372	0.9994520652827757	**7'10"623**	Desired	0,0001634422489649
26	**Multiscale variance gauss:**								
27	multiscale-fastluma-variance-gauss-zero	0.9922456777748451	10'46"450	**0.9925613635695231**	10'45"591	0.9921209715600335	10'46"792	Anomaly	0,0004403920094890
28	multiscale-fastluma-variance-gauss-constant	0.9969417840367856	**10'45"498**	**0.9973489384242628**	10'45"826	0.9971699357492918	**10'44"890**	Expected	0,0004071543874770
29	multiscale-mediumluma-variance-gauss-zero	**0.9999923347191011**	**1'48"638**	0.9999920183852874	**1'48"576**	0.9999913723334465	**1'48"514**	Desired	0,0000009623856551
30	multiscale-mediumluma-variance-gauss-constant	0.9922446054756333	**10'45"46**	**0.992560087763587**	**10'45"92**	0.992119448608418	**10'44"842**	Anomaly	0,0004406391551690
31	**Multiscale gradient unfiltered:**								
32	multiscale-fastluma-gradient-unfiltered-zero	**0.8346552466886107**	1'48"545	0.8341445419896498	1'48"405	0.8250869414632609	1'48"389	Desired	0,0095683052253499
33	multiscale-fastluma-gradient-unfiltered-constant	0.9884056477922687	**1'48"452**	**0.9895943501252445**	1'48"420	0.9888977581117036	1'48"467	Expected	0,0011887023329751
34	multiscale-mediumluma-gradient-unfiltered-zero	**0.8346544119371898**	1'49"247	0.8341435512167331	1'48"607	0.8250857567796904	**1'48"436**	Desired	0,0095686551575001
35	multiscale-mediumluma-gradient-unfiltered-constar	0.98840435914483	**1'48"654**	**0.9895928285808364**	**1'48"404**	0.988895963418122	**1'48"374**	Expected	0,0011884694360060
36	**Multiscale variance unfiltered:**								
37	multiscale-fastluma-variance-unfiltered-zero	**0.999994031815254**	1'48"404	0.9999939596400979	1'48"249	0.9999936034401464	1'47"999	Desired	0,0000004283751079
38	multiscale-fastluma-variance-unfiltered-constant	**0.9999936401328211**	1'48"529	0.9999935572885805	1'48"513	0.9999931880980742	1'48"248	Desired	0,0000004520347471
39	multiscale-mediumluma-variance-unfiltered-zero	**0.9996155075317397**	**7'10"873**	0.9995468035337404	**7'11"310**	0.9994520652827757	**7'10"967**	Desired	0,0001634422489649
40	multiscale-mediumluma-variance-unfiltered-constar	**0.9999929488867569**	**1'48"186**	0.9999926700827774	**1'47"827**	0.9999920639987804	1'48"280	Desired	0,0000008848879770

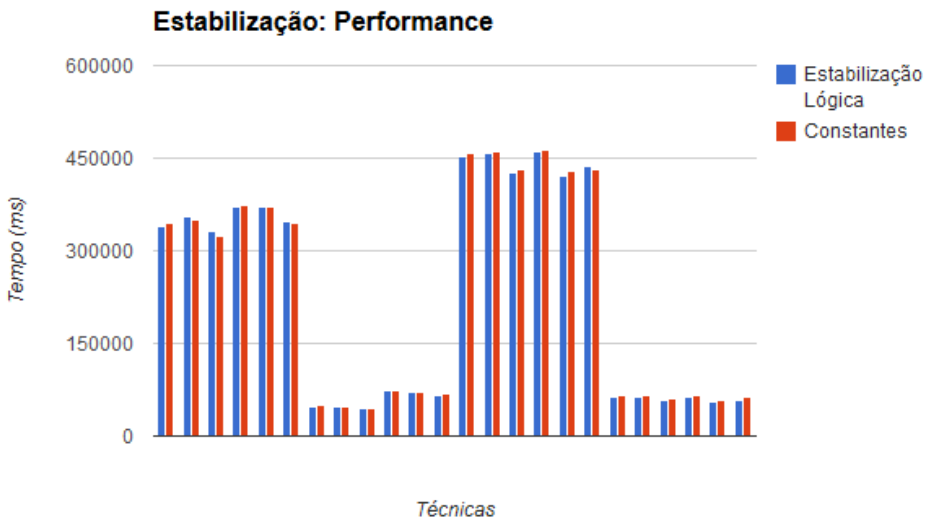
Resultados: Luminância



- Original vs Subamostrada
  - Mantém valores iguais até a 5ª decimal (0,00001)
  - Tempo +20ms/frame em escala original; em múltiplas escalas, ganho se dilui e perde importância

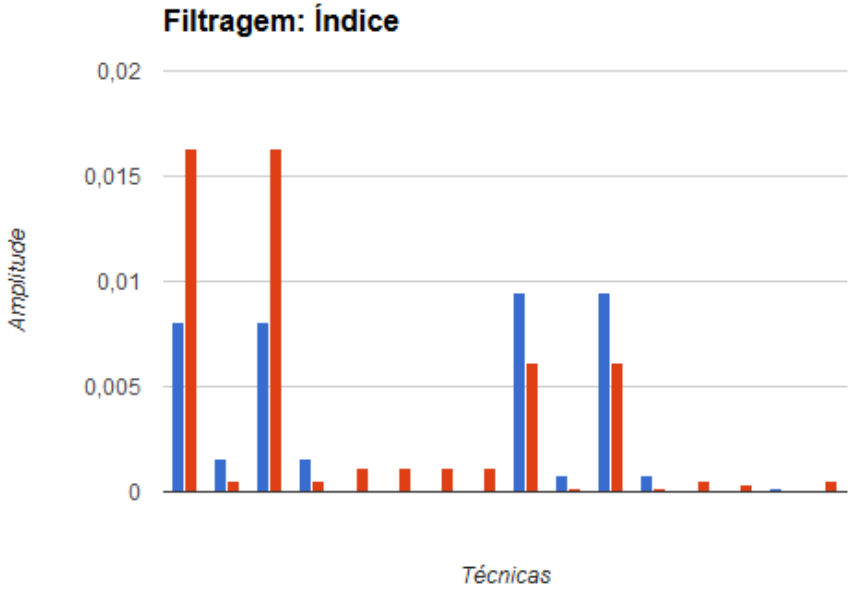
Resultados: Estabilização de Zeros

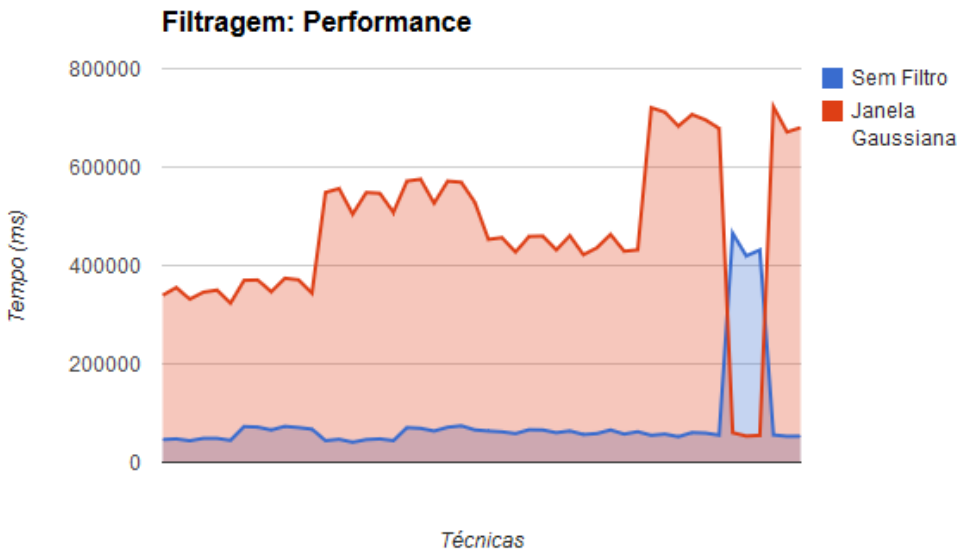




- Lógica vs. Constantes
  - Lógica aumenta amplitude de índice: útil para diferenciar resultados
  - Mesma performance
  - Bom para gradiente; variância inverte o ranking
  - Efeito atenuado em múltiplas escalas

Resultados: Janela de Filtragem





- Janela gaussiana vs. Sem filtro
  - Sem filtro melhora até 10x a performance
  - Na maioria dos casos, piora a amplitude de índice e embaralha o ranking
  - Com gradiente e estabilização lógica, respeita o ranking desejado e mantém amplitude (investigação futura)

Resultados: Diversos

- Ranking desejado foi atingido em 17 das 32 métricas.
- 2 métricas apresentam anomalia de performance (provável bug).
- Confirmação de que luminância é pouco relevante para todos os casos.
- Variância com Gauss é 50% mais lenta que gradiente com Gauss e produz resultados menos úteis.
- Gradiente apresenta mais amplitude que Variância; combinado com estabilização lógica, sempre adere ao ranking desejado; também adere no caso de estabilização por constante combinada a filtro de Gauss.
- A métrica ideal é: escala original + luminância subamostrada + estrutura gradiente + sem filtro + estabilizado logicamente.

Bibliografia

- WANG, Zhou; BOVIK, Alan C.; SHEIKH, H. R.; SIMONCELLI, E. P. **Image quality assessment**: From error visibility to structural similarity. IEEE Transactions on Image Processing, vol. 13, no. 4, p. 600-612, 2004. IEEE, abr. 2004.
- WANG, Zhou; SIMONCELLI, E. P.; BOVIK, Alan C. **Multi-scale structural similarity for image quality assessment**. In IEEE ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS. IEEE, nov. 2003.
- ROUSE, David; HEMAMI, Sheila. **Understanding and simplifying the structural similarity metric**. In IEEE International Conference in Image Processing 2008, San Diego, EUA.
- ROUSE, David; HEMAMI, Sheila. **Analyzing the role of visual structure in the recognition of natural image content with multi-scale SSIM**. In SPIE Human Vision and Electronic Imaging XIII, San José, EUA, 2008.
- CHEN, Guan-Hao; YANG, Chun-Ling; XIE, Sheng-Li. **Gradient-based structural similarity for image quality assessment**. In IEEE International Conference in Image Processing 2006, Atlanta, EUA.
- CHEN, Ming-Jun; BOVIK, Alan C. **Fast structural similarity index algorithm**. Journal of Real-Time Image Processing, vol. 6, ed. 4, p. 281-287. Dez. 2011.