Visual attention modelling and applications
Towards perceptual-based editing methods

Olivier Le Meur
olemeur@irisa.fr

IRISA - University of Rennes 1

November 18, 2014
Outline

1. Introduction
2. Visual attention
3. Image inpainting
4. Conclusion & Perspective
Visual attention
Inpainting
O. Le Meur

Introduction

Visual attention
Presentation
Computational models
Our cognitive model
Performance
Extension to video
Other contributions
Applications
Perspectives

Image inpainting
Examplar-based inpainting
Variants of Criminisi’s method
Some examples
Limitations
Super-resolution-based inpainting method
Results
Perspectives

Conclusion & Perspective
Three main topics: visual attention, image editing and compression
Visual Attention

2 Visual attention
- Presentation
- Computational models
- Our cognitive model
- Performance
- Extension to video
- Other contributions
- Applications
- Perspectives
Natural visual scenes are cluttered and contain many different objects that cannot all be processed simultaneously.

Amount of information coming down the optic nerve $10^8 - 10^9$ bits per second

Far exceeds what the brain is capable of processing...

Where is Waldo, the young boy wearing the red-striped shirt...
WE DO NOT SEE EVERYTHING AROUND US!!!

Visual attention

Posner proposed the following definition (Posner, 1980). Visual attention is used:

- to select important areas of our visual field (alerting);
- to search a target in cluttered scenes (searching).

There are several kinds of visual attention:

- **Overt visual attention**: involving eye movements;
- **Covert visual attention**: without eye movements (Covert fixations are not observable).
Bottom-Up vs Top-Down

- **Bottom-Up**: some things draw attention reflexively, in a task-independent way (Involuntary, Very quick, Unconscious);
- **Top-Down**: some things draw volitional attention, in a task-dependent way (Voluntary; Very slow; Conscious).
Bottom-Up vs Top-Down

- **Bottom-Up**: some things draw attention reflexively, in a task-independent way (Involuntary; Very quick; Unconscious);
- **Top-Down**: some things draw volitional attention, in a task-dependent way (Voluntary; Very slow; Conscious).

**Computational models of Bottom-up overt visual attention**
Most of the computation models of visual attention have been motivated by the seminal work of Koch and Ullmann (Koch and Ullman, 1985).

- a plausible computational architecture to predict our gaze;
- a set of feature maps processed in a massively parallel manner;
- a single topographic saliency map.
Taxonomy of models:

- Information Theoretic models;
- Cognitive models;
- Graphical models;
- Spectral analysis models;
- Pattern classification models;
- Bayesian models.

Extracted from (Borji and Itti, 2013).
Cognitive models:

- inspired by cognitive concepts;
- based on the HVS properties.

Extracted from (Borji and Itti, 2013).
In (Le Meur et al., 2006b), we designed a computational model of bottom-up visual attention.

1. **Input color image**;
2. Projection into a perceptual color space;
3. Subband decomposition in the Fourier domain;
4. CSF and Visual Masking;
5. Difference of Gaussians;
6. Pooling.
In (Le Meur et al., 2006b), we designed a computational model of bottom-up visual attention.

1. **Input color image**;
2. **Projection into a perceptual color space**;
3. Subband decomposition in the Fourier domain;
4. CSF and Visual Masking;
5. Difference of Gaussians;
6. Pooling.
In (Le Meur et al., 2006b), we designed a computational model of bottom-up visual attention.

1. **Input color image**;
2. **Projection into a perceptual color space**;
3. **Subband decomposition in the Fourier domain**;
4. **CSF and Visual Masking**;
5. **Difference of Gaussians**;
6. **Pooling**.
Our cognitive model (1/2)

In (Le Meur et al., 2006b), we designed a computational model of bottom-up visual attention.

1. Input color image;
2. Projection into a perceptual color space;
3. Subband decomposition in the Fourier domain;
4. CSF and Visual Masking;
5. Difference of Gaussians;
6. Pooling.
In (Le Meur et al., 2006b), we designed a computational model of bottom-up visual attention.

1. Input color image;
2. Projection into a perceptual color space;
3. Subband decomposition in the Fourier domain;
4. CSF and Visual Masking;
5. Difference of Gaussians;
6. Pooling.
In (Le Meur et al., 2006b), we designed a computational model of bottom-up visual attention.

1. Input color image;
2. Projection into a perceptual color space;
3. Subband decomposition in the Fourier domain;
4. CSF and Visual Masking;
5. Difference of Gaussians;
6. Pooling.
Our cognitive model (2/2)

Good prediction:

Failure cases:
Our cognitive model (2/2)

- Good prediction:

- Failure cases:
Performance on still images (1/3)

The requirement of a ground truth

- Eye tracker:
- A panel of observers;
- An appropriate protocol.

Adapted from (Judd et al., 2009).
Discrete fixation map \( f^i \) for the \( i^{th} \) observer:

\[
  f^i(x) = \sum_{k=1}^{M} \delta(x - x_k)
\]

where \( M \) is the number of fixations and \( x_k \) is the \( k^{th} \) fixation.

Continuous saliency map \( S \):

\[
  S(x) = \left( \frac{1}{N} \sum_{i=1}^{N} f^i(x) \right) \ast G_\sigma(x)
\]

where \( N \) is the number of observers.
Performance of the proposed model in terms of linear correlation, extracted from (Borji et al., 2012).

- more than 30 models;
- good performance of the proposed model, given that it dates back to 2006.
Our cognitive model for video sequences (1/2)

The temporal salience is the difference between local and dominant motion \(\text{(Le Meur et al., 2007)}\).

We estimate the dominant motion using a complete 2D affine motion model:

\[
\begin{align*}
\overrightarrow{V}_\Theta(s) &= \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \Theta = \begin{pmatrix} a_1 + a_2 x + a_3 y \\ a_4 + a_5 x + a_6 y \end{pmatrix}
\end{align*}
\]

where \(\Theta = [a_1, a_2, a_3, a_4, a_5, a_6]^T\) are the affine parameters of the model.

The problem is solved by using M-estimator \(\text{(Odobez and Bouthemy, 1995)}\):

\[
\hat{\Theta} = \arg\min_\Theta \sum_{s \in S} \rho(r(s))
\]

where \(r(s) = I(s + \overrightarrow{V}_\Theta(s), t + 1) - I(s, t)\) represents the displaced frame difference. \(\rho\) is a symmetric, positive-definite function.
Our cognitive model for video sequences (2/2)
Robustness of saliency models (Le Meur, 2011): Repetitiveness, i.e. the ability of saliency models to provide similar results in different impairment conditions, is high.
Robustness of saliency models (Le Meur, 2011):

Repeatability, i.e. the ability of saliency models to provide similar results in different impairment conditions, is high.

Time-dependent saliency map (Gautier and Le Meur, 2012):

\[ S(x, t) = \sum_{k=1}^{K} p_k(t) \phi_k(x) \]

where, \( \phi_k() \) represents a feature map and \( p_k \) its time-dependent weight (\( \sum_i p_i = 1, p_i \geq 0 \ \forall i \)).
Robustness of saliency models (Le Meur, 2011):

Repeatability, i.e. the ability of saliency models to provide similar results in different impairment conditions, is high.

Time-dependent saliency map (Gautier and Le Meur, 2012):

\[ S(x, t) = \sum_{k=1}^{K} p_k(t) \phi_k(x) \]

where, \( \phi_k() \) represents a feature map and \( p_k \) its time-dependent weight (\( \sum_i p_i = 1, p_i \geq 0 \forall i \)).

2D vs 3D (Khaustova et al., 2014):
Difference between 2D and 3D visual attention when the video content is displayed with crossed disparity.
Robustness of saliency models (Le Meur, 2011):

Repeatability, i.e. the ability of saliency models to provide similar results in different impairment conditions, is high.

Time-dependent saliency map (Gautier and Le Meur, 2012):

\[ S(x, t) = \sum_{k=1}^{K} p_k(t) \phi_k(x) \]

where, \( \phi_k() \) represents a feature map and \( p_k \) its time-dependent weight (\( \sum_i p_i = 1, p_i \geq 0 \ \forall i \)).

2D vs 3D (Khaustova et al., 2014):

Difference between 2D and 3D visual attention when the video content is displayed with crossed disparity.

Saliency aggregation (Le Meur and Liu, 2014):

Aggregation of existing saliency maps to improve the quality.
Applications (1/2)

Compression: Thanks to saliency maps, it might be possible to improve upon the quality of visually important areas.

Quality assessment (Ninassi et al., 2007, 2009): Artifacts on salient areas might be more annoying than those appearing on non-salient areas.

Retargeting (Chamaret et al., 2010, Le Meur et al., 2006a): "The intelligent reframing technology allows adapting any content to any screen format and size. It improves the reframing quality and the viewer experience."

Conclusion & Perspective
Applications (1/2)

Compression:

Thanks to saliency maps, it might be possible to improve upon the quality of visually important areas.
Applications (1/2)

→ **Compression:**

Thanks to saliency maps, it might be possible to improve upon the quality of visually important areas.

→ **Quality assessment (Ninassi et al., 2007, 2009):**

Artifacts on salient areas might be more annoying than those appearing on non salient areas.
Applications (1/2)

- **Compression:**
  
  Thanks to saliency maps, it might be possible to improve upon the quality of visually important areas.

- **Quality assessment (Ninassi et al., 2007, 2009):**
  
  Artifacts on salient areas might be more annoying than those appearing on non salient areas.

- **Retargeting (Chamaret et al., 2010, Le Meur et al., 2006a):**
  
  “The intelligent reframing technology allows adapting any content to any screen format and size. It improves the reframing quality and the viewer experience.”
Memorability (Mancas and Le Meur, 2013):

The memorability of an image could be predicted by using saliency-based low-level features.

Compared to (Isola et al., 2011):

- improvement of 2% in terms of correlation coefficient;
- 512 features have been replaced by 17 saliency-based features for the learning.
Applications (2/2)

- **Memorability** *(Mancas and Le Meur, 2013):*

  The memorability of an image could be predicted by using saliency-based low-level features.

  Compared to *(Isola et al., 2011):*
  - improvement of 2% in terms of correlation coefficient;
  - 512 features have been replaced by 17 saliency-based features for the learning.

- **Salient object detection** *(Liu et al., 2014):*

  The goal is to extract salient objects from complex scenes.
The picture is much clearer than 10 years ago!

BUT...

➤ How far are we from human performance?*
➤ How to take advantage of top-down influences?*
➤ What is the best score?*
➤ What is the most representative dataset?*

✗ Important aspects of our visual system are clearly overlooked;
✗ Current models implicitly assume that eyes are equally likely to move in any direction;
✗ Systematic biases are not taken into account;
✗ The temporal dimension is not considered (static saliency map)...

* Open challenges of visual attention, tutorial CVPR’13, Borji et al.
Saccadic model to infer the saliency map

Let be an image \( I : \Omega \subset \mathbb{R}^n \leftrightarrow \mathbb{R}^m \), the goal is to predict the visual scanpath (composed of a set of fixations \( x \)).

\[
x_t^* = \arg \max_{x \in \Omega} p(x|x_{t-1}, \ldots, x_{t-T})
\]

\[
p(x|x_{t-1}, \ldots, x_{t-T}) \propto p_{BU}(x)p_B(d, \phi)p_M(x, t)
\]

\( p_{BU} \) would be the bottom-up saliency map;
\( p_B \) would be the systematic tendencies of oculomotor behavior;
\( p_M \) would be the memory (Inhibition-of-Return);
Be able to predict the fixation duration.
Short saccades around 1 to 3 degrees of visual angle.
Systematic tendencies on natural scenes

- Anisotropic shape;
- More horizontal saccades than vertical ones;
- Very few diagonal saccades.

Distribution of saccade orientation
Systematic tendencies on natural scenes

- Joint distribution of saccade amplitudes and orientations;
- Strong horizontal bias and small saccade.
Perspectives: visual attention modelling (4/4)

Systematic tendencies on webpage

- Joint distribution of saccade amplitudes and orientations;
- Strong horizontal bias in the rightward direction;
- F-shaped pattern.

From (Shen and Zhao, 2014)'s eye fixation dataset.
Let be an image \( I \) defined as
\[
I : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m
\]
with \( \Omega = S \cup U \),
- \( S \) being the known part of \( I \)
- \( U \) the unknown part of \( I \)

Let be a degradation operator \( M \)
\[
M : \Omega \rightarrow \{0, 1\}
\]
\[
M(x) = \begin{cases} 
0, & \text{if } x \in U \\
1, & \text{otherwise}
\end{cases}
\]

Let \( F \) the observed image: \( F = M \circ I \)
Let be an image $I$ defined as

$$I : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$$

with $\Omega = S \cup U$,

- $S$ being the known part of $I$
- $U$ the unknown part of $I$

Let be a degradation operator $M$

$$M : \Omega \rightarrow \{0, 1\}$$

$$M(x) = \begin{cases} 0, & \text{if } x \in U \\ 1, & \text{otherwise} \end{cases}$$

Let $F$ the observed image: $F = M \circ I$
Examplar-based inpainting methods rely on the assumption that the known part of the image provides a good dictionary which could be used efficiently to restore the unknown part (Efros and Leung, 1999).

The recovered texture is therefore learned from similar regions.

- This can be done simply by sampling, copying or combining patches from the known part of the image; **Template Matching**
- Patches are then stitched together to fill in the missing area.
Examplar-based inpainting (2/3)

Criminisi et al.’s algorithm (Criminisi et al., 2004)

Criminisi et al. has brought a new momentum to inpainting applications and methods. They proposed a new method based on two sequential stages:

1. Filling order computation
2. Texture synthesis by recopy

Notations:

- A patch $\psi_{px}$ is a discretized $N \times N$ neighborhood centered on the pixel $px$. This patch can be vectorized in a raster-scan order as a $mN^2$-dimensional vector;

- $\psi_{px}^{uk}$ denotes the unknown pixels of the patch;

- $\psi_{px}^k$ denotes its known pixels;

- $\psi_{px}(i)$ denotes the $i^{th}$ nearest neighbour of $\psi_{px}$;

- $\delta U$ is the front line;
Filling order computation: $P(p_x) = C(p_x) \times D(p_x)$

Confidence term

$$C(p_x) = \frac{1}{|\psi_{p_x}|} \sum_{q \in \psi_{p_x}^k} C(q)$$

where $|\psi_{p_x}|$ is the area of $\psi_{p_x}$.

Data term

$$D(p_x) = \frac{\left| \nabla I^\perp(p_x) \cdot n_{p_x} \right|}{\alpha}$$

where $\alpha$ is a normalization constant.
Examplar-based inpainting (3/3)

1. Filling order computation: \( P(p_x) = C(p_x) \times D(p_x) \)

   **Confidence term**
   
   \[
   C(p_x) = \frac{1}{|\psi_{p_x}|} \sum_{q \in \psi^k_{p_x}} C(q)
   \]

   **Data term**
   
   \[
   D(p_x) = \frac{|\nabla I(p_x) \cdot n_{p_x}|}{\alpha}
   \]

   where \( \alpha \) is a normalization constant.

2. Texture synthesis by recopy:

   A template matching is performed within a local neighborhood:

   \[
   p_y = \arg \min_{q \in \mathcal{W}} d(\psi^k_{p_q}, \psi^k_{p_x^*})
   \]

   \( \mathcal{W} \subseteq S \) is the window search;

   \( \psi^k_{p_x^*} \) are the known pixels of the patch \( \psi_{p_x^*} \) with the highest priority;

   \( d(a, b) \) is the sum of squared differences between patches \( a \) and \( b \).

The pixels of the patch \( \psi^u_{p_y} \) are then copied into \( \psi^u_{p_x^*} \).
Instead of using the gradient, (Le Meur et al., 2011) used the structure tensor which is more robust:

\[ D(p_x) = \alpha + (1 - \alpha) \exp \left( -\frac{\eta}{(\lambda_1 - \lambda_2)^2} \right) \]

where \( \eta \) is a positive value and \( \alpha \in [0, 1] \).

The structure tensor is a symmetric, positive semi-definite matrix (Weickert, 1999):

\[ J_{\rho,\sigma}[I] = K_\rho * \left( \sum_{i=1}^{m} \nabla (I_i * K_\sigma) \nabla (I_i * K_\sigma)^T \right) \]

where \( K_\alpha \) is a Gaussian kernel with a standard deviation \( \alpha \). The parameters \( \rho \) and \( \sigma \) are called integration scale and noise scale, respectively.
Filling order computation (2/2)

\[ D(p_x) = \alpha + (1 - \alpha) \exp \left( -\frac{\eta}{(\lambda_1 - \lambda_2)^2} \right) \]

When \( \lambda_1 \simeq \lambda_2 \), the data term tends to \( \alpha \). It tends to 1 when \( \lambda_1 \gg \lambda_2 \).
Texture synthesis with more than one candidate

From $K$ patches $\psi_{p_x(i)}$ which are the most similar to the known part $\psi_{p_x}^k$ of the input patch, the unknown part of the patch to be filled $\widehat{\psi}_{p_x}^{uk}$ is then obtained by a linear combination of the sub-patches $\psi_{p_x(i)}^{uk}$.

$$\widehat{\psi}_{p_x}^{uk} = \sum_{i=1}^{K} w_i \psi_{p_x(i)}^{uk}$$

How can we compute the weights $w_i$ of this linear combination?

Note: $K$ is locally adjusted by using an $\epsilon$-ball including patches within a certain radius.
Texture synthesis (2/2)

\[ \hat{\psi}_{p_x}^{uk} = \sum_{i=1}^{K} w_i \psi_{p_x(i)}^{uk} \]

Different solutions exist \((\text{Guillemot et al., 2013})\):

- Average template matching: \( w_i = \frac{1}{K}, \forall i \);
- Non-local means approach \((\text{Buades et al., 2005})\):
  \[ w_i = \exp \left( -\frac{d(\psi_{p_x}^k, \psi_{p_x(i)}^k)}{h^2} \right) \]
- Least-square method minimizing \((\text{Turkan and Guillemot, 2012})\)
  \[ E(w) = \| \psi_{p_x}^k - Aw \|_2^2, a \]

- with or without the constraint: the weights sum to one;
- with or without the constraint: the weights are positive.

The crucial role of the similarity metric...
Results from (Le Meur et al., 2011).
Limitations (1/1)

- **Very sensitive to the parameter settings** such as the filling order and the patch size:

- **Examplar-based methods** are one-pass greedy algorithms.
Objectives of the proposed method (Le Meur and Guillemot, 2012, Le Meur et al., 2013)

We apply an examplar-based inpainting algorithm several times and fuse together the inpainted results.

- less sensitive to the inpainting setting;
- relax the greedy constraint.
Objectives of the proposed method (Le Meur and Guillemot, 2012, Le Meur et al., 2013)

We apply an examplar-based inpainting algorithm several times and fuse together the inpainted results.

- less sensitive to the inpainting setting;
- relax the greedy constraint.

The inpainting method is applied on a coarse version of the input picture:

- less demanding of computational resources;
- less sensitive to noise;
- $K$ candidates for the texture synthesis without introducing blur.
Objectives of the proposed method (Le Meur and Guillemot, 2012, Le Meur et al., 2013)

We apply an examplar-based inpainting algorithm several times and fuse together the inpainted results.

- less sensitive to the inpainting setting;
- relax the greedy constraint.

The inpainting method is applied on a coarse version of the input picture:

- less demanding of computational resources;
- less sensitive to noise;
- \( K \) candidates for the texture synthesis without introducing blur.

Need to fuse the inpainted images and to retrieve the highest frequencies

Loopy Belief Propagation and Super-Resolution algorithms.
The baseline algorithm is an examplar-based method:

- Filling order computation;
- Texture synthesis.

### Decimation factor $n = 3$

- 13 sets of parameters
More than one inpainting (1/1)

The baseline algorithm is an examplar-based method:
- **Filling order computation**;
- **Texture synthesis**.

- Decimation factor $n = 3$
- 13 sets of parameters

### Table: Thirteen inpainting configurations.

<table>
<thead>
<tr>
<th>Setting</th>
<th>Parameters</th>
</tr>
</thead>
</table>
| 1       | Patch’s size $5 \times 5$
|         | Decimation factor $n = 3$
|         | Search window $80 \times 80$
|         | Sparsity-based filling order |
| 2       | default + rotation by 180 degrees |
| 3       | default + patch’s size $7 \times 7$
| 4       | default + rotation by 180 degrees
|         | + patch’s size $7 \times 7$
| 5       | default + patch’s size $11 \times 11$
| 6       | default + rotation by 180 degrees
|         | + patch’s size $11 \times 11$
| 7       | default + patch’s size $9 \times 9$
| 8       | default + rotation by 180 degrees
|         | + patch’s size $9 \times 9$
| 9       | default + patch’s size $9 \times 9$
|         | + Tensor-based filling order |
| 10      | default + patch’s size $7 \times 7$
|         | + Tensor-based filling order |
| 11      | default + patch’s size $5 \times 5$
|         | + Tensor-based filling order |
| 12      | default + patch’s size $11 \times 11$
|         | + Tensor-based filling order |
| 13      | default + rotation by 180 degrees
|         | + patch’s size $9 \times 9$
|         | + Tensor-based filling order |
Loopy Belief Propagation (1/3)

Let be a finite set of labels \( \mathbb{L} \) composed of \( M = 13 \) values.

\[
E(l) = \sum_{p \in \nu} V_d(l_p) + \lambda \sum_{(n,m) \in N_4} V_s(l_n, l_m)
\]

where, \( l_p \) the label of pixel \( p_x \), \( \nu \) represents the pixel in \( U \) and \( N_4 \) is a neighbourhood system. \( \lambda \) is a weighting factor.

- \( V_d(l_p) \) represents the cost of assigning a label \( l_p \) to a pixel \( p_x \);
- \( V_s(l_n, l_m) \) is the discontinuity cost.

The minimization is performed iteratively (less than 15 iterations) (Boykov and Kolmogorov, 2004, Boykov et al., 2001, Yedidia et al., 2005).
LBP convergence:

- 13 inpainted image in input;
- 25 iterations;
- resolution = 80 x 120.
Loopy Belief Propagation (3/3)

LBP convergence:

- 13 inpainted image in input;
- 25 iterations;
- resolution=120 × 80.
For the LR patch corresponding to the HR patch having the highest priority:

- We look for its best neighbor;
- Only the best candidate is kept;
- The corresponding HR patch is simply deduced.
- Its pixel values are then copied into the unknown parts of the current HR patch.
Results (1/4)
Results (2/4)
Results (3/4)
Results (4/4)
 Perspectives: inpainting (1/2)

- Understanding the scene layout to improve on the inpainting quality (structure, depth...);

- Would it be possible to evaluate the quality of an inpainted image?

- Revisiting texture synthesis:
  - Texture mixing for synthesizing a new texture from a collection of examplars.
  - Generating *perceptually* similar texture rather than copying?
  - Optimal mass transportation problem.
Perspectives: inpainting (2/2)

- Extension to deal with video inpainting.
  Video are available on author’s webpage.

- Extension to multiview, HDR, RGB-D...
Conclusion (1/2)

Research topics:

- Visual attention modelling;
- Image inpainting;
- Super-resolution (Ferreira et al., 2014);
- HEVC compression (Dhollande et al., 2014);
- Examplar-based style transfer.

Spring to autumn...
Conclusion (1/2)

Research topics:

- Visual attention modelling;
- Image inpainting;
- Super-resolution (Ferreira et al., 2014);
- HEVC compression (Dhollande et al., 2014);
- Examplar-based style transfer.

Spring to autumn...
Research topics:

- Visual attention modelling;
- Image inpainting;
- Super-resolution (Ferreira et al., 2014);
- HEVC compression (Dhollande et al., 2014);
- Examplar-based style transfer.

Spring to autumn...

Bibliometrics

- Book chapters: 3
- Conference: 34 (ACCV, ECCV, ICIP, HVEI, ICME, ACM MM...)
Perspectives (2/2)

Visual Attention:

- Visual scanpath modelling;
- The use of high-level information;
- Better datasets, better scoring metrics;
- Computational medicine...
Perspectives (2/2)

Visual Attention:
- Visual scanpath modelling;
- The use of high-level information;
- Better datasets, better scoring metrics;
- Computational medicine...

Image Inpainting:
- Video inpainting (HDR, RGB-D...);
- Quality metric...

Visual attention
Inpainting
O. Le Meur
Introduction
Visual attention
Presentation
Computational models
Our cognitive model
Performance
Extension to video
Other contributions
Applications
Perspectives

Image inpainting
Examplar-based inpainting
Variants of Criminisi’s method
Some examples
Limitations
Super-resolution-based inpainting method
Results
Perspectives

Conclusion & Perspective
Perspectives (2/2)

**Visual Attention:**
- Visual scanpath modelling;
- The use of high-level information;
- Better datasets, better scoring metrics;
- Computational medicine...

**Image Inpainting:**
- Video inpainting (HDR, RGB-D...);
- Quality metric...

---

**Perceptual-based image editing**

How to modify an image to maximize its saliency, its memorability?

How *to compel* observers to look unconsciously where we want...
Perspectives (2/2)

Visual Attention:
- Visual scanpath modelling;
- The use of high-level information;
- Better datasets, better scoring metrics;
- Computational medicine...

Image Inpainting:
- Video inpainting (HDR, RGB-D...);
- Quality metric...

Perceptual-based image editing
How to modify an image to maximize its saliency, its memorability?
How to compel observers to look unconsciously where we want...

inpainting, super-resolution, gradient-boosting, scanpath-based metric...
Thanks for your attention
Un grand MERCI!

Doctorants:
Alexandre Ninassi, Brice Follet, Josselin Gautier, Mounira Ebdelli, Darya Khaustova.

Etudiants de master:
Olivier Gaborieau, Guillaume Courtin, Ayodeji Aribuki, Younesse Andam, David Wolinski, Alan Bourasseau, Julien Sicre, Hristina Hristova.

A mes anciens collègues de GrassValley et de Technicolor
A mes encadrants de thèse
A l’équipe SIROCCO et plus particulièrement à Christine.

A ma famille!!
A Karen, Glen, Louen et Anouk
References


Visual attention

Inpainting

O. Le Meur

References


Chengyao Shen and Qi Zhao. Webpage saliency. In ECCV. IEEE, 2014.

