Introduction to Information Theory

Tutorial on Information Theory in Visualization

Mateu Sbert

University of Girona, Spain

Tianjin University, China

EUROGRAPHICS 2016

Overview

- Introduction
- Information measures
  - entropy, conditional entropy
  - mutual information
- Information channel
- Relative entropy
- Mutual information decomposition
- Inequalities
- Information bottleneck method
- Entropy rate
- Continuous channel

Introduction (1)

- Claude Elwood Shannon, 1916-2001
- "A mathematical theory of communication", Bell System Technical Journal, July and October, 1948
- The significance of Shannon's work
- Transmission, storage and processing of information
- Applications: physics, computer science, mathematics, statistics, biology, linguistics, neurology, computer vision, etc.

Introduction (2)

- Certain quantities, like entropy and mutual information, arise as the answers to fundamental questions in communication theory
- Shannon entropy is the ultimate data compression or the expected length of an optimal code
- Mutual information is the communication rate in presence of noise

- Book: T.M. Cover and J.A. Thomas, Elements of Information Theory, Wiley, 1991, 2006

Introduction (3)

- Shannon introduced two fundamental concepts about "information" from the communication point of view
  - information is uncertainty
    - information source is modeled as a random variable or a random process
    - probability is employed to develop the information theory
  - information to be transmitted is digital
    - Shannon's work contains the first published use of "bit"

- Book: R.W. Yeung, Information Theory and Network , Springer, 2008

Information Measures (1)

- Random variable X taking values in an alphabet $\mathcal{X}$

$$\mathcal{X}:\{x_1,x_2,\ldots,x_n\},\ p(x)=\Pr\{X=x\},\ p(X)=\{p(x),x\in\mathcal{X}\}$$

- Shannon entropy $H(X)$, $H(p)$: uncertainty, information, homogeneity, uniformity

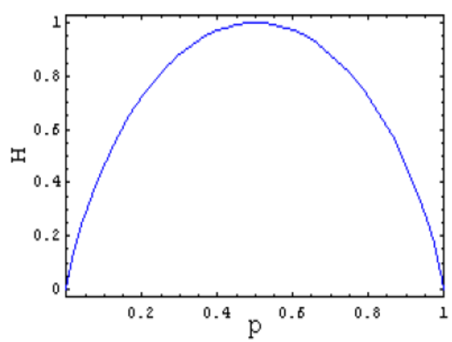
$$H(X)=-\sum_{x\in\mathcal{X}}p(x)\log p(x)\equiv-\sum_{i=1}^{n}p(x_i)\log p(x_i)$$

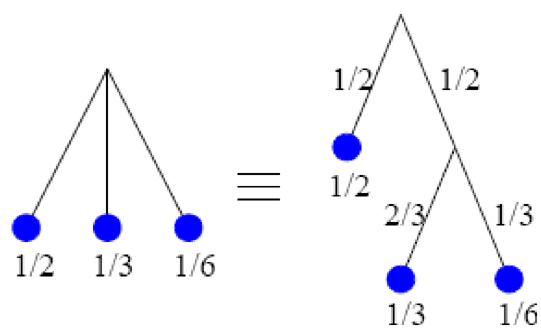
  - information associated with x: $-\log p(x)$; base of logarithm: 2; convention: $0\log 0=0$; unit: bit: uncertainty of the toss of an ordinary coin

Information Measures (2)

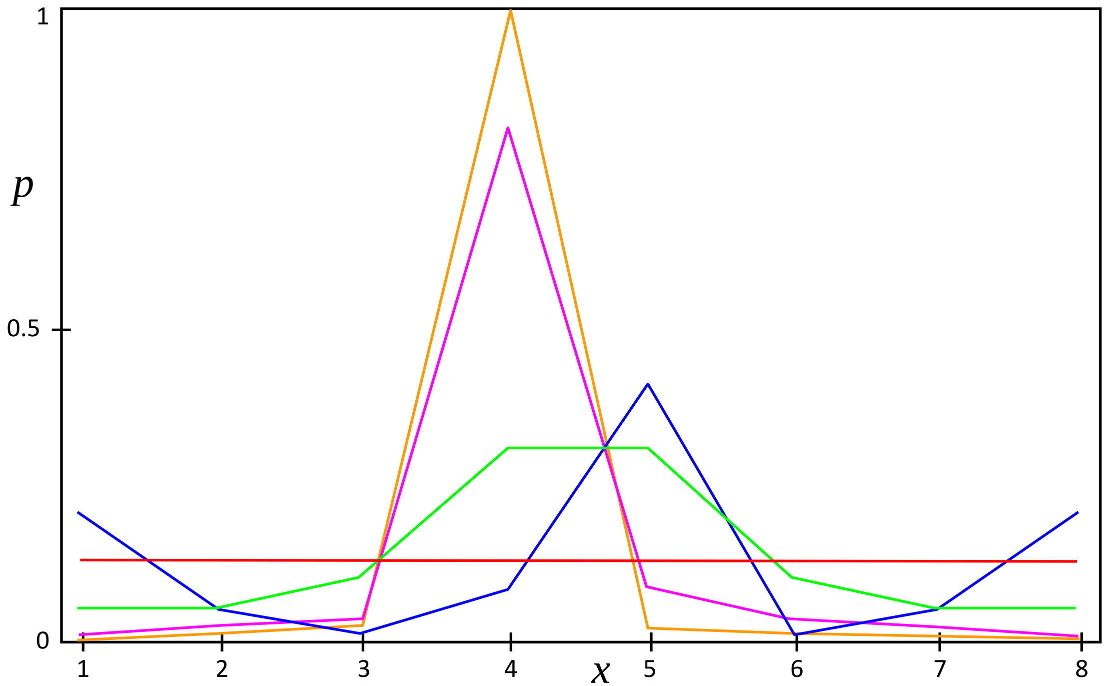
For example, the entropy of a fair coin toss is $H(X) = -(1/2)\log(1/2) - (1/2)\log(1/2) = \log 2 = 1$ bit. For the toss of a fair die with alphabet $\mathcal{X} = \{1, 2, 3, 4, 5, 6\}$ and probability distribution $p(X) = \{1/6, 1/6, 1/6, 1/6, 1/6, 1/6\}$, the entropy is $H(X) = \log 6 = 2.58$ bits.

- Properties of Shannon entropy
  - $0 \le H(X) \le \log|X|$
  - binary entropy: $H(X) = -p\log p - (1-p)\log(1-p)$
  - $H(p_1, p_2, p_3) = H(p_1, p_2 + p_3) + (p_2 + p_3) H\left(\frac{p_2}{p_2 + p_3}, \frac{p_3}{p_2 + p_3}\right)$





Information Measures (3)



H(0.001, 0.002, 0.003, 0.980, 0.008, 0.003, 0.002, 0.001) = 0.190
H(0.010, 0.020, 0.030, 0.800, 0.080, 0.030, 0.020, 0.010) = 1.211
H(0.200, 0.050, 0.010, 0.080, 0.400, 0.010, 0.050, 0.200) = 2.314
H(0.125, 0.125, 0.125, 0.125, 0.125, 0.125, 0.125, 0.125) = 3.000

Information Measures (4)

- Discrete random variable Y in an alphabet $\mathcal{Y}$

$$\mathcal{Y}: \{y_1, y_2, \ldots, y_n\}, p(y) = \Pr\{Y = y\}$$

- Joint entropy $H(X,Y)$

$$H(X,Y) = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log p(x,y)$$

- Conditional entropy $H(Y|X)$

$$H(Y|X) = \sum_{x \in \mathcal{X}} p(x) H(Y|x) = -\sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y|x) \log p(y|x)$$

$$= -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log p(y|x)$$

Information Channel

- Communication or information channel $X \to Y$

$$\underset{p(X)}{X} \xrightarrow{p(Y|X)} \underset{p(Y)}{Y}$$

$p(X)$ $\begin{pmatrix} p(x_1) \\ p(x_2) \\ \dots \\ p(x_n) \end{pmatrix}$

$p(Y|X)$ $\begin{pmatrix} p(y_1|x_1) & p(y_2|x_1) & \dots & p(y_m|x_1) \\ p(y_1|x_2) & p(y_2|x_2) & \dots & p(y_m|x_2) \\ \dots & \dots & \dots & \dots \\ p(y_1|x_n) & p(y_2|x_n) & \dots & p(y_m|x_n) \end{pmatrix}$ $p(Y|x)$

$p(Y)$ $\begin{pmatrix} p(y_1) & p(y_2) & \dots & p(y_m) \end{pmatrix}$

$$\sum_{x \in X} p(y|x) = 1$$

$$p(y) = \sum_{x \in X} p(x) p(y|x)$$

Bayes' rule $p(x,y) = p(x)p(y|x) = p(y)p(x|y)$

Information Measures (5)

- Mutual information *I(X;Y)*: shared information, correlation, dependence, information transfer

$$I(X;Y) = H(Y) - H(Y|X) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

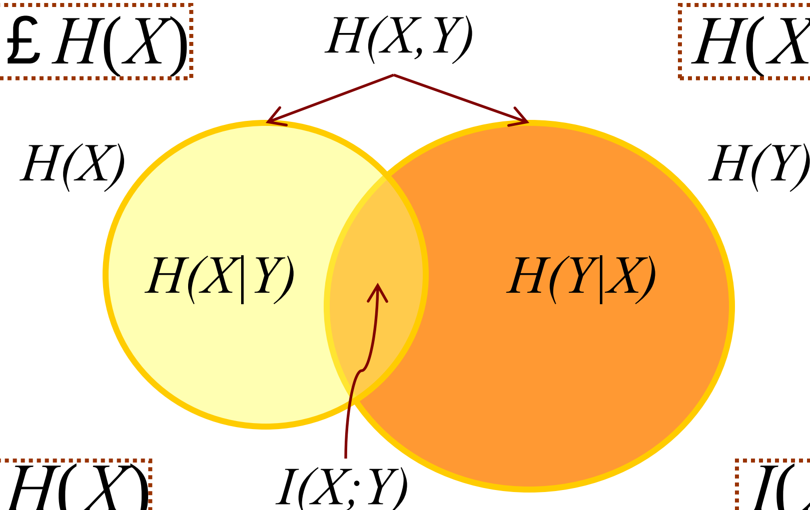
$$= \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log \frac{p(y|x)}{p(y)}$$

Information Measures (6)

- Relationship between information measures

$$0 \le H(X|Y) \le H(X)$$

$$H(X,Y) = H(X) + H(Y|X)$$



$$I(X;Y) \le H(X)$$

$$I(X;Y) = I(Y;X) \ge 0$$

$$H(X,Y) = H(X) + H(Y) - I(X;Y)$$

Yeung's book: Chapter 3 establishes a one-to-one correspondence between Shannon's information measures and set theory. A number of examples are given to show how the use of information diagrams can simplify the proofs of many results in information theory.

Information Measures (7)

As an example, we consider the joint distribution $p(X,Y)$ represented in Fig. 1.3.*left*. The marginal probability distributions of X and Y are given by $p(X) = \{0.25, 0.25, 0.5\}$ and $p(Y) = \{0.375, 0.625\}$, respectively. Thus, $H(X) = -0.25 \log 0.25 - 0.25 \log 0.25 - 0.5 \log 0.5 = 1.5$ bits, $H(Y) = -0.375 \log 0.375 - 0.625 \log 0.625 = 0.954$ bits, and $H(X,Y) = -0.125 \log 0.125 - 0.125 \log 0.125 - 0.25 \log 0.25 - 0 \log 0 - 0 \log 0 - 0.5 \log 0.5 = 1.75$ bits.

$p(X,Y)$		$\mathcal{Y}$: y_1	y_2	$p(X)$
$\mathcal{X}$	x_1	0.125	0.125	0.25
	x_2	0.25	0	0.25
	x_3	0	0.5	0.5
$p(Y)$		0.375	0.625	
$H(X,Y) = 1.75$				

$p(Y\|X)$		$\mathcal{Y}$: y_1	y_2	$H(Y\|x \in \mathcal{X})$
$\mathcal{X}$	x_1	0.5	0.5	$H(Y\|x_1) = 1$
	x_2	1	0	$H(Y\|x_2) = 0$
	x_3	0	1	$H(Y\|x_3) = 0$
$H(Y\|X) = 0.25$				

$$
\begin{aligned}
H(Y|X) &= \sum_{i=1}^{3} p(x_i) H(Y|X = x_i) \\
&= 0.25\ H(Y|X = x_1) + 0.25\ H(Y|X = x_2) + 0.5\ H(Y|X = x_3) \\
&= 0.25 \times 1 + 0.25 \times 0 + 0.5 \times 0 = 0.25 \text{ bits.}
\end{aligned}
$$

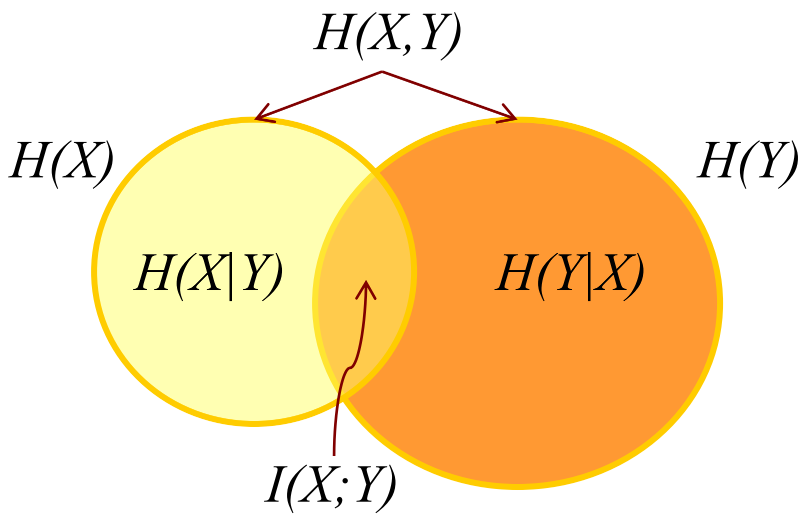
Information Measures (8)

- Normalized mutual information: different forms

$$\frac{I(X;Y)}{H(X,Y)} \qquad \frac{I(X;Y)}{H(X)+H(Y)}$$

$$\frac{I(X;Y)}{\max\{H(X),H(Y)\}}$$

$$\frac{I(X;Y)}{\min\{H(X),H(Y)\}}$$



- Information distance

$$H(X\,|\,Y)+H(Y\,|\,X)$$

Relative Entropy

- Relative entropy, informational divergence, Kullback-Leibler distance $D_{KL}(p,q)$: how much p is different from q (on a common alphabet $\mathcal{X}$)

$$D_{KL}(p,q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}$$

  - convention: $0 \log 0/q = 0$ and $p \log p/0 = \infty$
  - $D_{KL}(p,q) >= 0$
  - it is not a true metric or "distance" (non-symmetric, triangular inequality is not fulfilled)
  - $I(X;Y) = D_{KL}(p(X,Y), p(X)p(Y))$

Mutual Information

$$I(X;Y) = H(Y) - H(Y|X) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

$$= \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log \frac{p(y|x)}{p(y)}$$

$$D_{KL}(p,q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}$$

$$I(X;Y) = D_{KL}(p(X,Y), p(X)p(Y))$$

Mutual Information Decomposition

- Information associated with x

$$I(X;Y) = \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log \frac{p(y|x)}{p(y)} = \sum_{x \in X} p(x)\left(H(Y) - H(Y|x)\right)$$

$$I_1(x;Y) = \sum_{y \in Y} p(y|x) \log \frac{p(y|x)}{p(y)}$$

$$I_2(x;Y) = H(Y) - H(Y|x)$$

[DeWeese]

$$I_3(x;Y) = \sum_{y \in Y} p(y|x) I_2(X;y)$$

[Butts]

$$I(X;Y) = \sum_{x \in X} p(x) I_k(x;Y)$$

$$k = 1, 2, 3$$

Mutual Information Decomposition

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - \sum_{x \in X} p(x) H(Y|x) = \sum_{x \in X} p(x)(H(Y) - H(Y|x))$$

$$= \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} = \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log \frac{p(y|x)}{p(y)}$$

$$I_1(x;Y) = \sum_{y \in Y} p(y|x) \log \frac{p(y|x)}{p(y)}$$

$$I_2(x;Y) = H(Y) - H(Y|x)$$

Inequalities

- **Data processing inequality**: if $X \to Y \to Z$ is a Markov chain, then

  $$I(X;Y) \geq I(X;Z)$$

  No processing of Y can increase the information that Y contains about X, i.e., further processing of Y can only increase our uncertainty about X on average
- **Jensen's inequality**: a function $f(x)$ is said to be convex over an interval (a,b) if for every x_1, x_2 in (a,b) and $0 \leq \lambda \leq 1$

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

Jensen-Shannon Divergence

- From the concavity of entropy, Jensen-Shannon divergence

$$JS(\pi_1,\ldots,\pi_N;p_1,\ldots,p_N)=H\left(\sum_{i=1}^{N}\pi_i p_i\right)-\sum_{i=1}^{N}\pi_i H(p_i)\geq 0$$

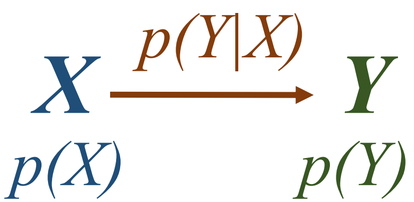
[Burbea]

- $$JS(\pi_1,\ldots,\pi_N;p_1,\ldots,p_N)=\sum_{i=1}^{N}\pi_i D_{KL}\left(p_i,\sum_{i=1}^{N}\pi_i p_i\right)$$

- $$JS(p(x_1),\ldots,p(x_n);p(Y|x_1),\ldots,p(Y|x_n))=I(X;Y)$$

Information Channel, MI and JS

- Communication or information channel $X \rightarrow Y$



$$p(X) = \begin{pmatrix} p(x_1) \\ p(x_2) \\ \dots \\ p(x_n) \end{pmatrix} \qquad p(Y|X) = \begin{pmatrix} p(y_1 | x_1) & p(y_2 | x_1) & \dots & p(y_m | x_1) \\ p(y_1 | x_2) & p(y_2 | x_2) & \dots & p(y_m | x_2) \\ \dots & \dots & \dots & \dots \\ p(y_1 | x_n) & p(y_2 | x_n) & \dots & p(y_m | x_n) \end{pmatrix}$$

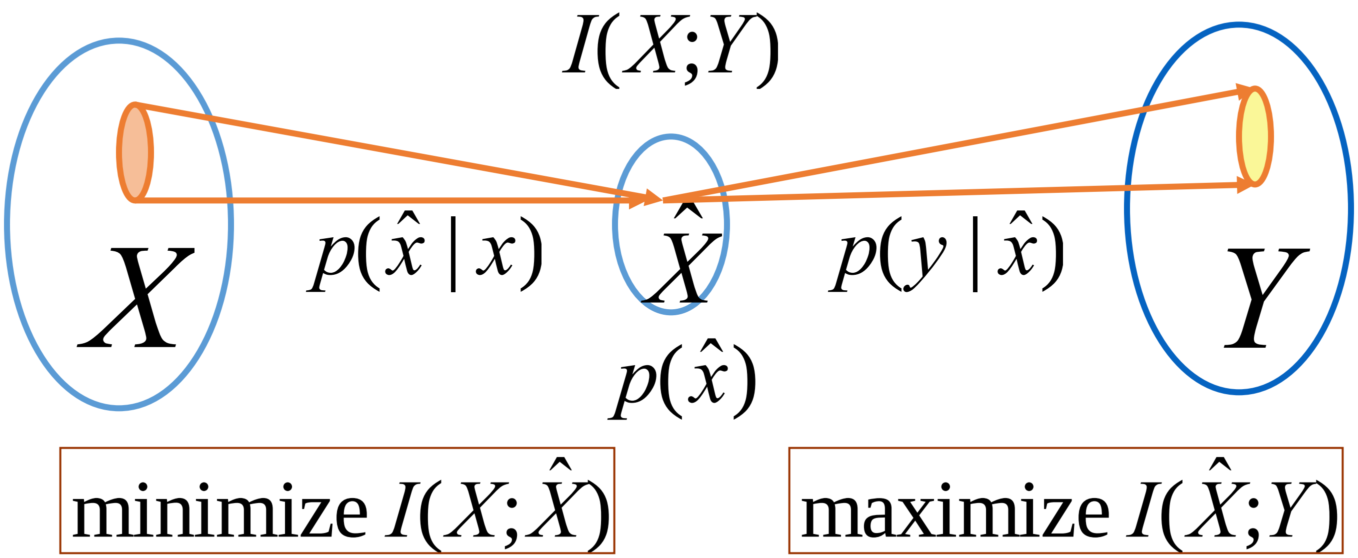
Each row: $p(Y|x)$

$$p(Y) = \begin{pmatrix} p(y_1) & p(y_2) & \dots & p(y_m) \end{pmatrix}$$

$$JS(p(x_1),\dots,p(x_n); p(Y|x_1),\dots,p(Y|x_n)) = I(X;Y)$$

Information Bottleneck Method (1)

- Tishby, Pereira and Bialek, 1999
- To look for a compressed representation of X which maintains the (mutual) information about the relevant variable Y as high as possible



Information Bottleneck Method (2)

- **Agglomerative information bottleneck method**: clustering/merging is guided by the minimization of the loss of mutual information

$$I(X;Y) \geq I(\hat{X};Y)$$

- Loss of mutual information

$$I(X;Y) - I(\hat{X};Y) = p(\hat{x}) JS(p(x_1)/p(\hat{x}),\ldots,p(x_m)/p(\hat{x}); p(Y|x_1),\ldots,p(Y|x_m))$$

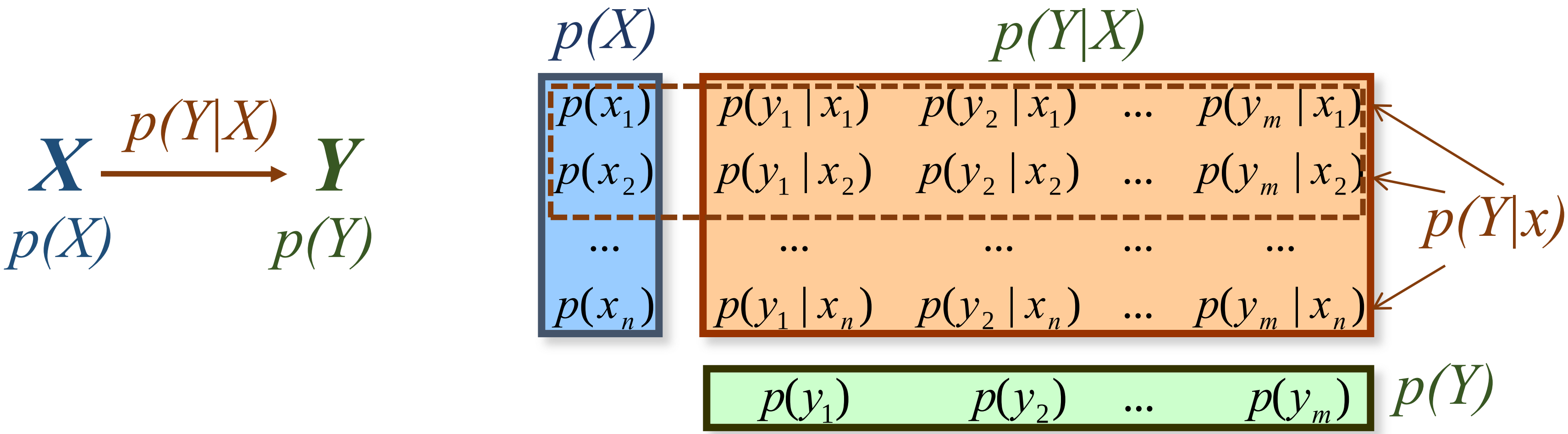
[Slonim]

where $p(\hat{x}) = \sum_{k=1}^{m} p(x_k)$

- The quality of each cluster $\hat{x}$ is measured by the Jensen-Shannon divergence between the individual distributions in the cluster

Information Channel and IB

- Communication or information channel $X \rightarrow Y$

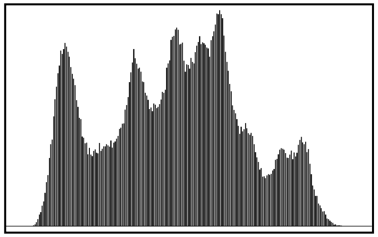


$$I(X;Y) - I(\hat{X};Y) =$$
$$p(\hat{x})JS(p(x_1)/p(\hat{x}), p(x_2)/p(\hat{x}); p(Y \mid x_1), p(Y \mid x_2))$$

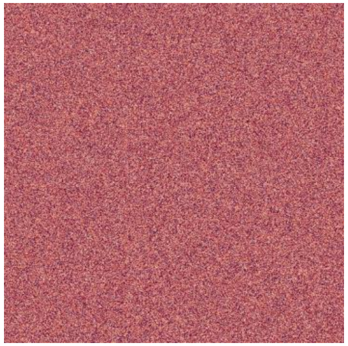
$$p(\hat{x}) = p(x_1) + p(x_2)$$

Example: Entropy of an Image

- The information content of an image is expressed by the Shannon entropy of the (normalized) intensity histogram



- The entropy disregards the spatial contribution of pixels



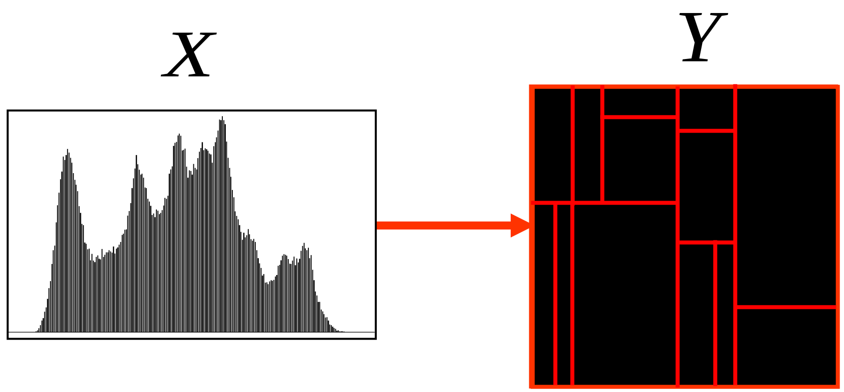
Example: Image Partitioning (1)

- Information channel $X \rightarrow Y$ defined between the intensity histogram and the image regions

$$X \xrightarrow{p(Y|X)} Y$$

$$p(X) \qquad\qquad p(Y)$$

$$p(X) = \left\{ \frac{b_i}{N} \right\}$$



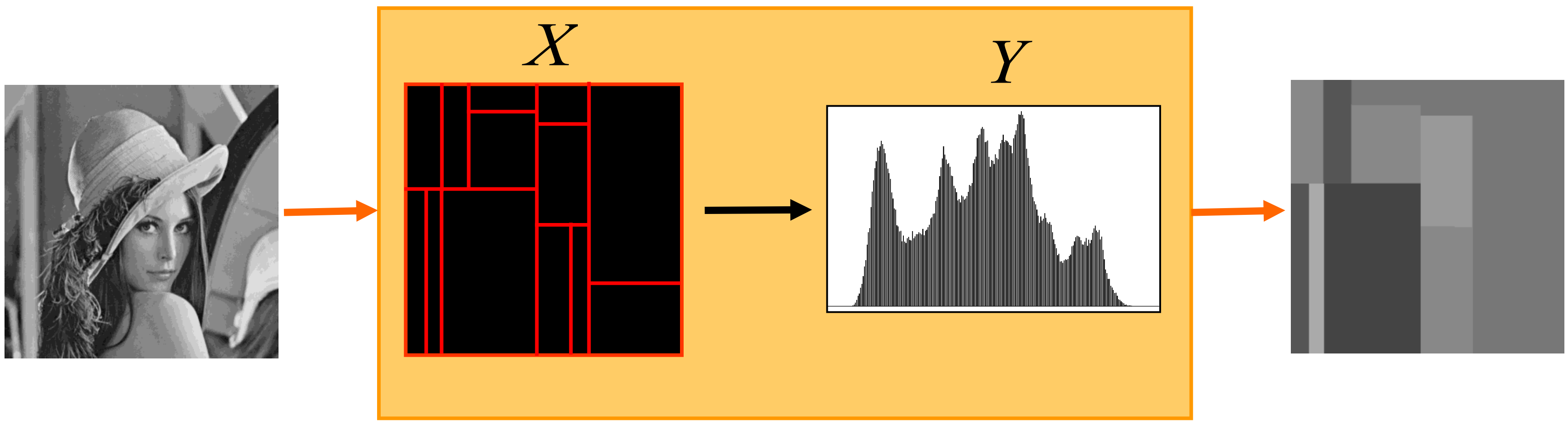
$$p(Y) = \left\{ \frac{r_j}{N} \right\}$$

b_i = number of pixels of bin i; r_j = number of pixels of region j

N = total number of pixels

Example: Image Partitioning (2)

information bottleneck method



information gain

$$I(X;Y) - I(\hat{X};Y) = p(\hat{x}) JS(p(x_1)/p(\hat{x}), p(x_2)/p(\hat{x}); p(Y|x_1), p(Y|x_2))$$

at each step, increase of $I(X;Y)$ = decrease of $H(X|Y)$

$$H(X) = I(X;Y) + H(X|Y)$$

Example: Image Partitioning (3)

$$MIR = \frac{I(\hat{X};Y)}{I(X;Y)}$$; number of regions ; % of regions

0.1; ***13***; *0.00*

0.2; ***64;*** *0.02*

0.3; ***330***; *0.13*

0.4; ***1553;*** *0.59*

0.0; ***5597***; *2.14*

1; ***234238***; *89.35*

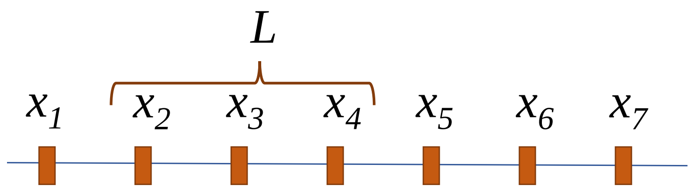
0.9; ***129136***; *49.26*

0.8; ***67291***; *25.67*

0.7; ***34011***; *12.97*

0.6; ***15316***; *5.84*

Entropy Rate



- Shannon entropy

$$H(X) = -\sum_{x\in\mathcal{X}} p(x)\log p(x)$$

- Joint entropy

$$H(X^L) = -\sum_{x^L\in\mathcal{X}^L} p(x^L)\log p(x^L)$$

- Entropy rate or information density

$$h = \lim_{L\to\infty} \frac{H(X^L)}{L}$$

$$= \lim_{L\to\infty} (H(X^L) - H(X^{L-1}))$$

Continuous Channel

- Continuous entropy

$$H^c(X) = -\int_S p(x)\log p(x)dx$$

$$\lim_{\Delta\to 0} H(X^\Delta) \neq H^c(X)$$

- Continuous mutual information

$$I^c(X,Y) = \int_S\int_S p(x,y)\log\frac{p(x,y)}{p(x)p(y)}dxdy$$

$$\lim_{\Delta\to 0} I(X^\Delta, Y^\Delta) = I^c(X,Y)$$

  - $I^c(X,Y)$ is the least upper bound for $I(X,Y)$
  - refinement can never decrease $I(X,Y)$

Viewpoint metrics and applications

Tutorial on Information Theory in Visualization

Mateu Sbert

University of Girona, Spain

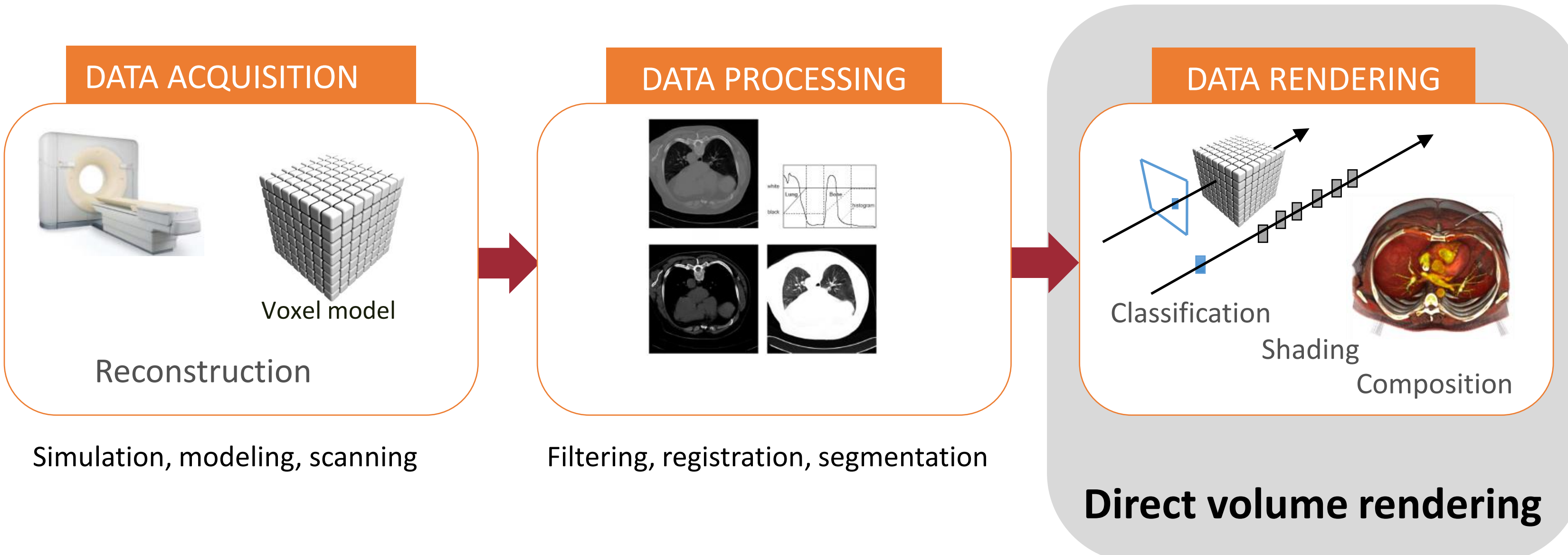
Tianjin University, China

EUROGRAPHICS 2016

Viewpoint selection

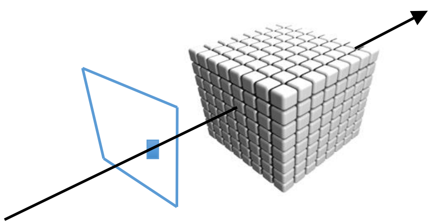
- Automatic selection of the most informative viewpoints is a very useful focusing mechanism in visualization
- It can guide the viewer to the most interesting information of the scene or data set
- A selection of most informative viewpoints can be used for a virtual walkthrough or a compact representation of the information the data contains
- Best view selection algorithms have been applied to computer graphics domains, such as scene understanding and virtual exploration, N best views selection , image-based modeling and rendering, mesh simplication, molecular visualization, and camera placement
- Information theory measures have been used as viewpoint metrics since the work of Vazquez et al. [2001], see also [Sbert et al. 2009]

The visualization pipeline



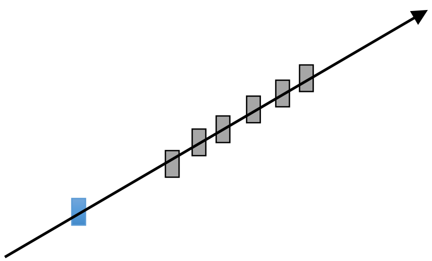
Direct volume rendering (DVR)

- Volume dataset is considered as a transparent gel with light travelling through it



- classification maps primitives to graphical attributes

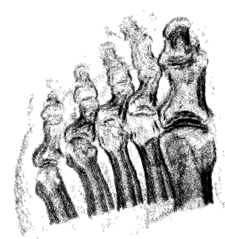
Transfer function definition



- shading (illumination) models shadows, light scattering, absorption…
  - usually absorption + emission optical model

Local or global illumination

- compositing integrates samples with optical properties along viewing rays



Both realistic and illustrative rendering

Viewpoint selection

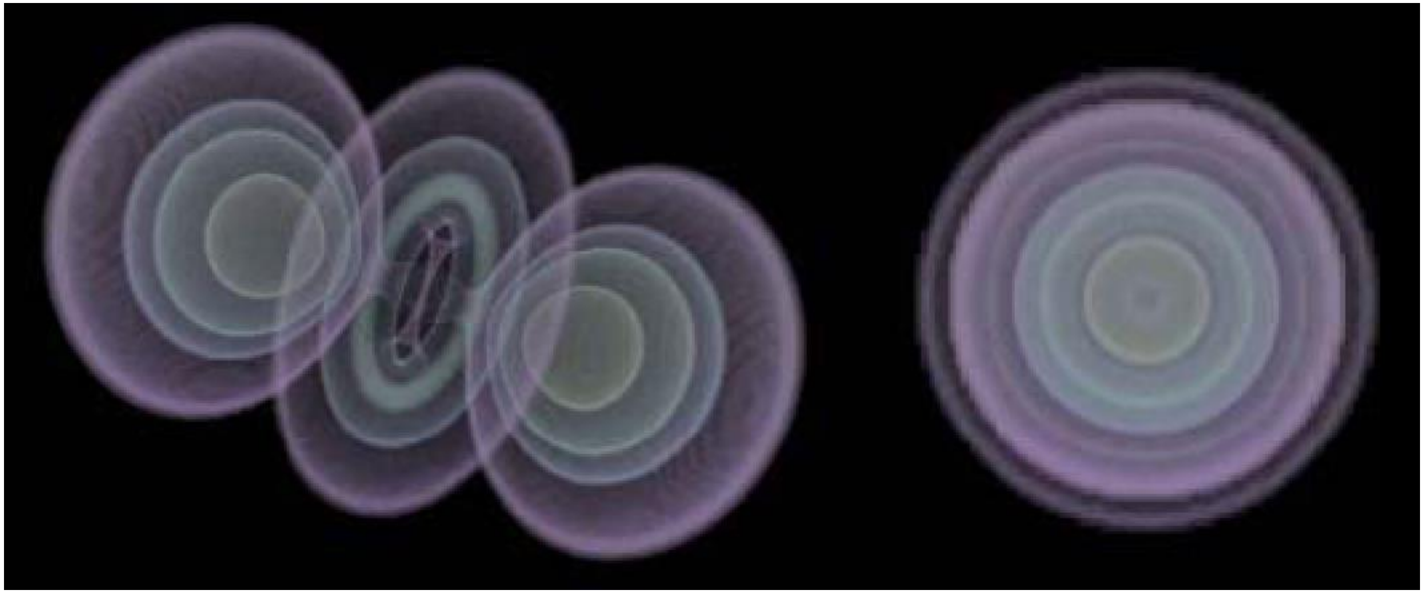
- Takahashi 2005
  - Evaluation of viewpoint quality based on the visibility of extracted isosurfaces or interval volumes.
  - Use as viewpoint metrics the average of viewpoint entropies for the extracted isosurfaces.

$$E_i^{iso}(v) = \frac{-1}{\log(m_i + 1)} \sum_{j=0}^{m_i} \frac{A_{ij}}{S} \log \frac{A_{ij}}{S}$$

$$E^{iso}(v) = \sum_{i=1}^{n} \frac{E_i^{iso}(v)}{n}$$

Viewpoint selection

- Takahashi et al.2005



Best and worst views of interval volumes extracted from a data set containing simulated electron density distribution in a hydrogen atom

Viewpoint selection

- Bordoloi and Shen 2005
  - Best view selection: use entropy of the projected visibilities distribution

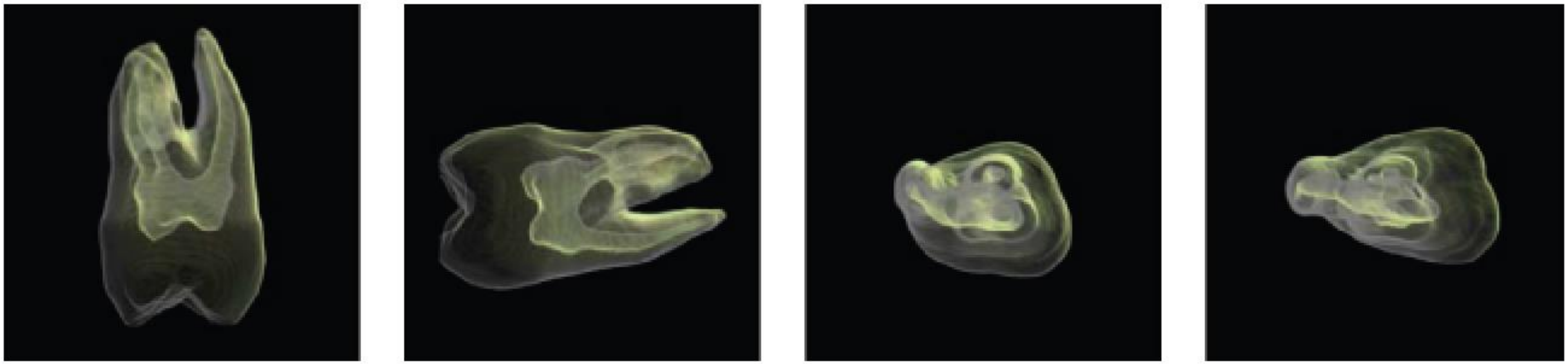
$$H(v) = -\sum_{i=1}^{n} q_i(v) \log q_i(v)$$

  - Representative views: cluster views according to Jensen-Shannon similarity measure

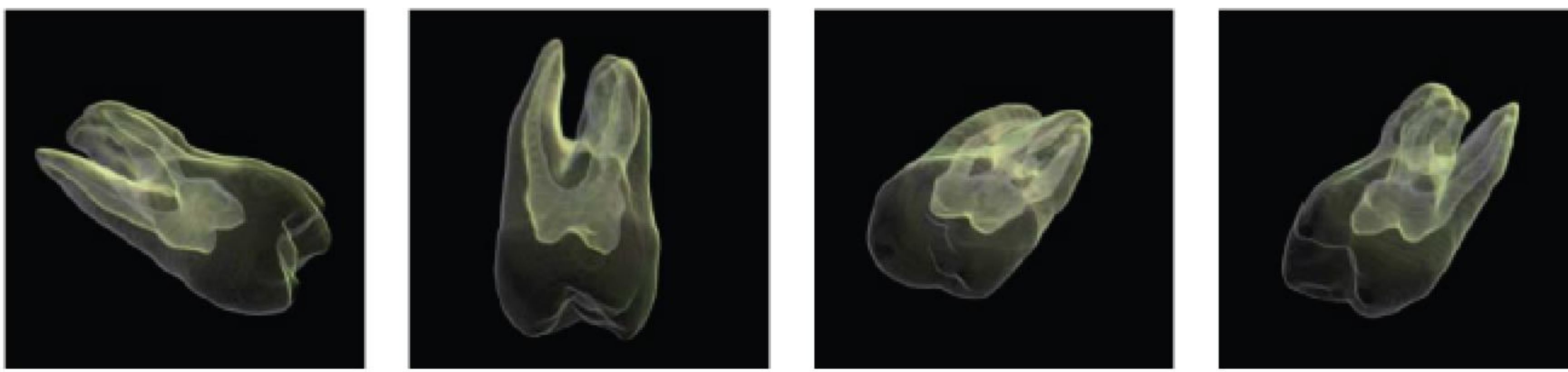
$$JS(\frac{1}{2}, \frac{1}{2}; q(v_1), q(v_2)) = H(\frac{1}{2}q(v_1) + \frac{1}{2}q(v_2)) - H(q(v_1)) - H(q(v_2))$$

Viewpoint selection

- Bordoloi and Shen 2005



Best (two left) and worst (two right) views of tooth data set

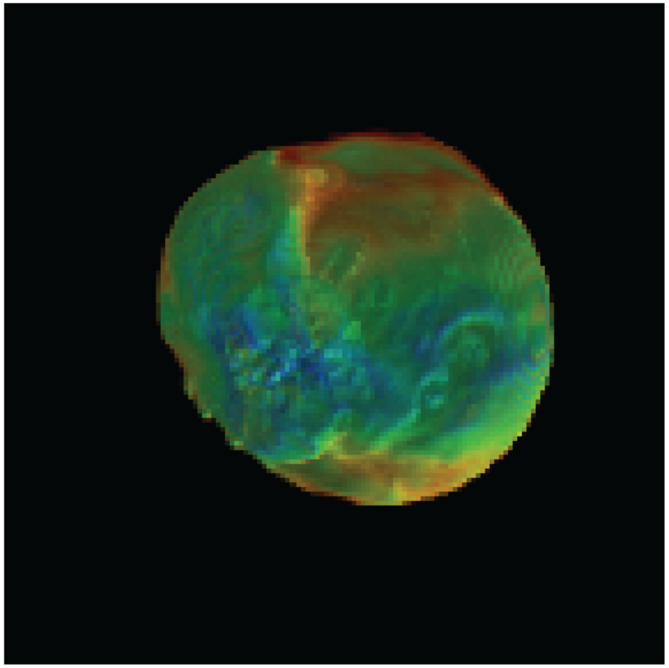


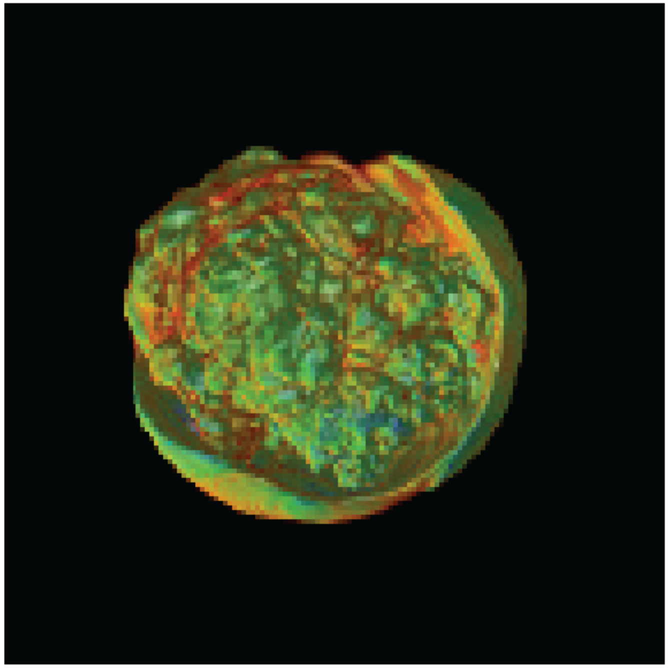
Four representative views

Viewpoint selection

- Ji and Shen 2006
  - Quality of viewpoint *v*, *u(v)*, is a combination of three values

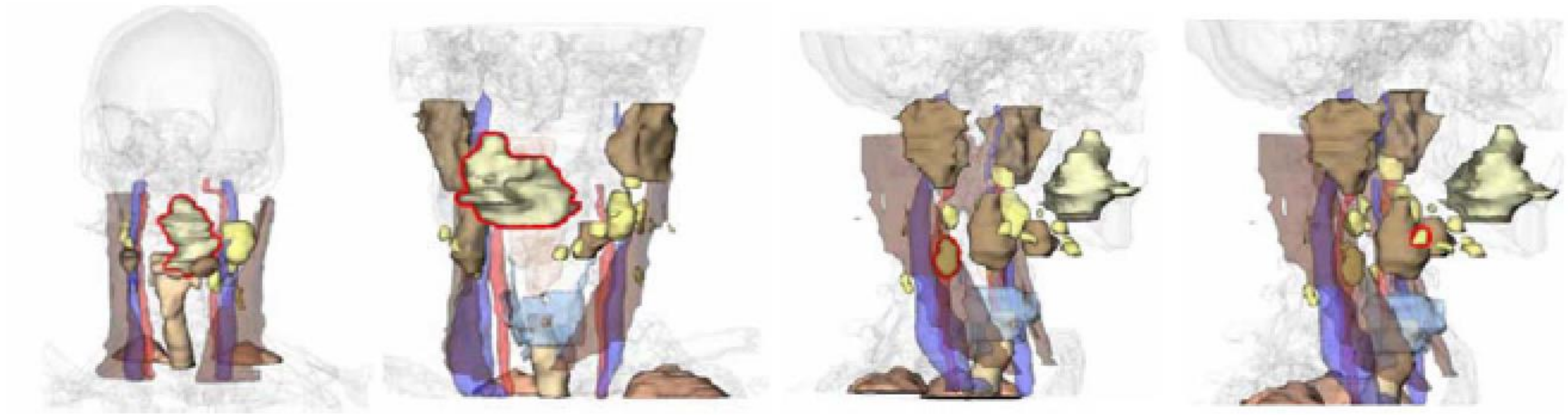
$$u(v) = \beta_1 opacity(v) + \beta_2 color(v) + \beta_3 curvature(v)$$





Viewpoint selection

- Mühler et al. 2007
  - Semantics-driven view selection. Entropy, between other factors, used to select best views.
  - Guided navigation through features assists studying the correspondence between focus objects.



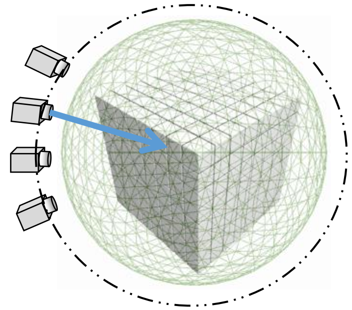
Visibility channel

- Viola et al. 2006, Ruiz et al. 2010

$$p(v) = \frac{vis(v)}{\sum_{i\in\mathcal{V}} vis(i)} \qquad p(z|v) = \frac{vis(z|v)}{vis(v)}$$

V $\xrightarrow{p(Z|V)}$ **Z**

$p(V)$ viewpoints — $p(Z)$ voxels



$p(V)$	$p(Z\|V)$			
$p(v_1)$	$p(z_1\|v_1)$	$p(z_2\|v_1)$	$\cdots$	$p(z_m\|v_1)$
$p(v_2)$	$p(z_1\|v_2)$	$p(z_2\|v_2)$	$\cdots$	$p(z_m\|v_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$p(v_n)$	$p(z_1\|v_n)$	$p(z_2\|v_n)$	$\cdots$	$p(z_m\|v_n)$
$p(Z)$	$p(z_1)$	$p(z_2)$	$\cdots$	$p(z_m)$

$$p(z) = \sum_{v\in\mathcal{V}} p(v)p(z|v)$$

- How a viewpoint sees the voxels
- Mutual information

$$I(V;Z) = \sum_{v\in\mathcal{V}} p(v) \sum_{z\in\mathcal{Z}} p(z|v)\log\frac{p(z|v)}{p(z)} = \sum_{v\in\mathcal{V}} p(v) I(v;Z)$$

- Viewpoint mutual information (VMI)

$$I(v;Z) = \sum_{z\in\mathcal{Z}} p(z|v)\log\frac{p(z|v)}{p(z)}$$

Reversed visibility channel

- Ruiz et al. 2010

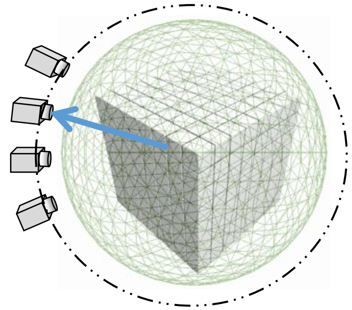
$$p(v|z) = \frac{p(v)p(z|v)}{p(z)}$$

$p(V|Z)$

Z → **V**

$p(Z)$ $p(V)$

voxels viewpoints



$p(Z)$	$p(V\|Z)$			
$p(z_1)$	$p(v_1\|z_1)$	$p(v_2\|z_1)$	$\cdots$	$p(v_m\|z_1)$
$p(z_2)$	$p(v_1\|z_2)$	$p(v_2\|z_2)$	$\cdots$	$p(v_m\|z_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$p(z_n)$	$p(v_1\|z_n)$	$p(v_2\|z_n)$	$\cdots$	$p(v_m\|z_n)$
$p(V)$	$p(v_1)$	$p(v_2)$	$\cdots$	$p(v_m)$

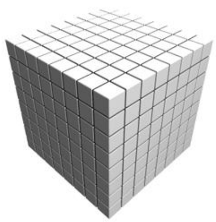
- How a voxel "sees" the viewpoints
- Mutual information

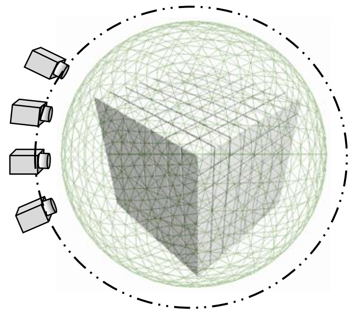
$$I(Z;V) = \sum_{z\in\mathcal{Z}} p(z) \sum_{v\in\mathcal{V}} p(v|z) \log \frac{p(v|z)}{p(v)} = \sum_{z\in\mathcal{Z}} p(z) I(z;V)$$

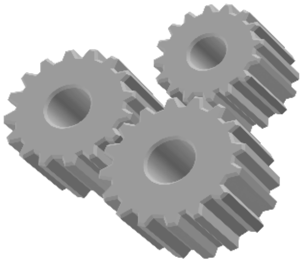
- Voxel mutual information (VOMI)

$$I(z;V) = \sum_{v\in\mathcal{V}} p(v|z) \log \frac{p(v|z)}{p(v)}$$

VOMI map computation

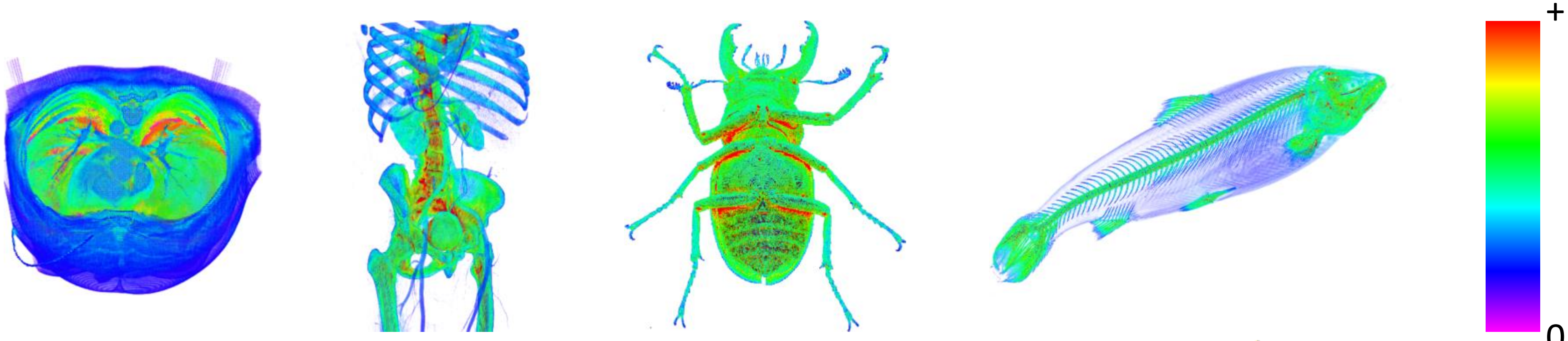






Transfer function

Volume dataset → Classified data → Ray casting → Visibility histogram for each viewpoint → Probabilities computation → **VOMI map**

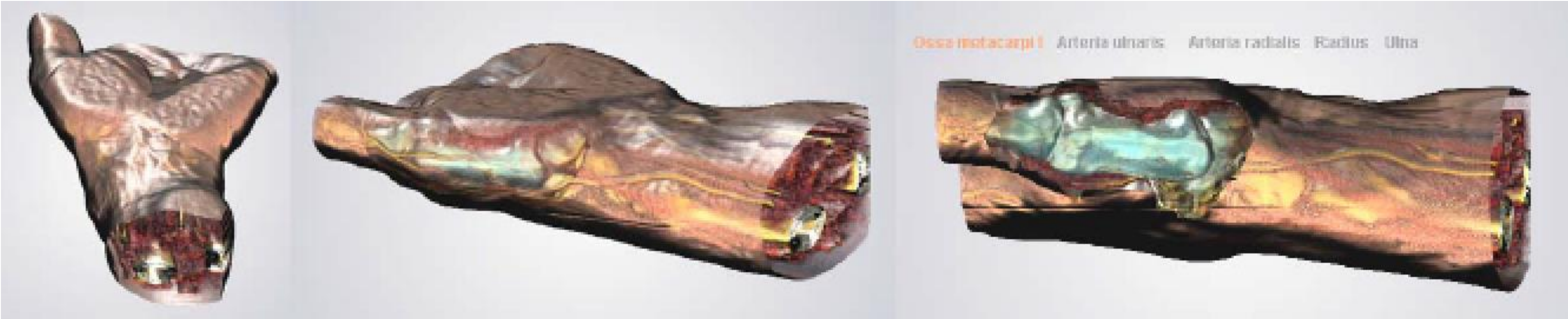


Visibility channel

- Viola et al. 2006
- Adding importance to VMI for viewpoint navigation with focus of interest. Objects instead of voxels

$$I(v;O) = D_{KL}(p(O|v)||p(O))$$

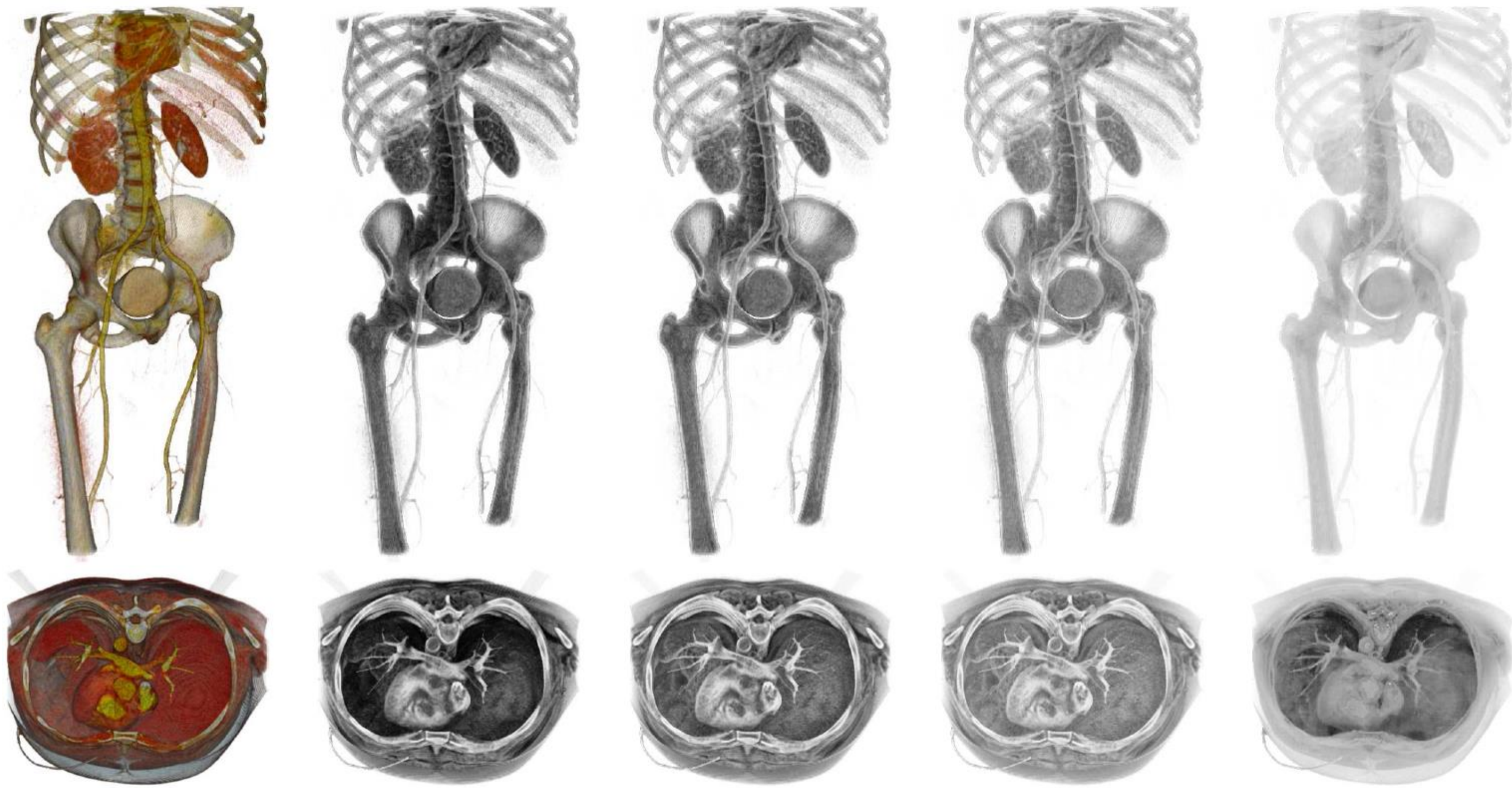
$$I'(v;O) = D_{KL}(p(O|v)||p'(O)) = \sum_{o \in \mathbb{O}} p(o|v) \log \frac{p(o|v)}{p'(o)}$$



VOMI applications

- Interpret VOMI as ambient occlusion
  - $AO(z) = 1 - \overline{I(z;V)}$
  - Simulate global illumination
  - Realistic and illustrative rendering
  - Color ambient occlusion
  - $CAO_{\alpha}(z;V) = \sum_{v \in \mathcal{V}} \left(p(v|z) \log \frac{p(v|z)}{p(z)} \right) \left(1 - C_{\alpha}(v)\right)$
- Interpret VOMI as importance
  - Modulate opacity to obtain focus+context effects emphasizing important parts
- "Project" VOMI to viewpoints to obtain informativeness of each viewpoint
  - $INF(v) = \sum_{z \in \mathcal{Z}} p(v|z) I(z;V)$
  - Viewpoint selection

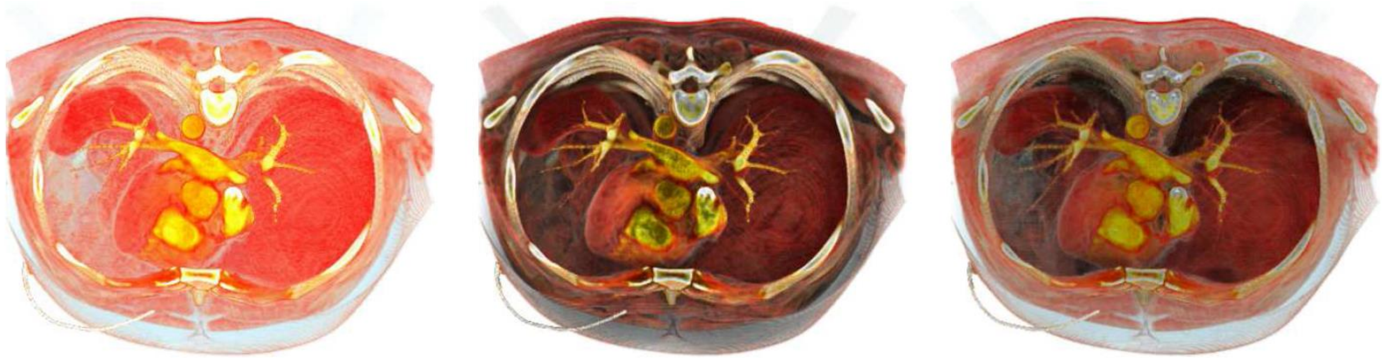
VOMI as ambient occlusion map



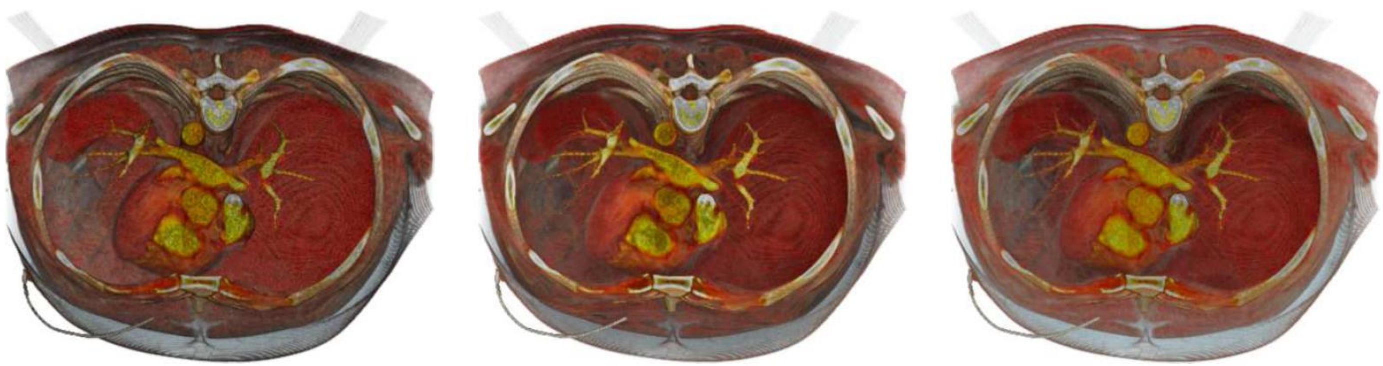
Original | Ambient Occlusion, Landis 2002 | Vicinity shading, Stewart 2003 | Obscurances, Iones et al. 98 | **VOMI**

VOMI applied as ambient occlusion

- Ambient lighting term

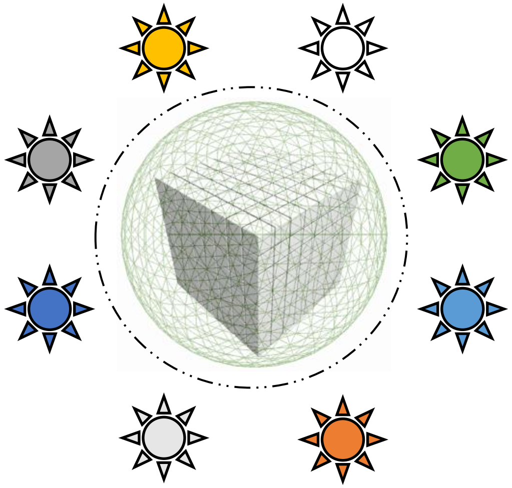


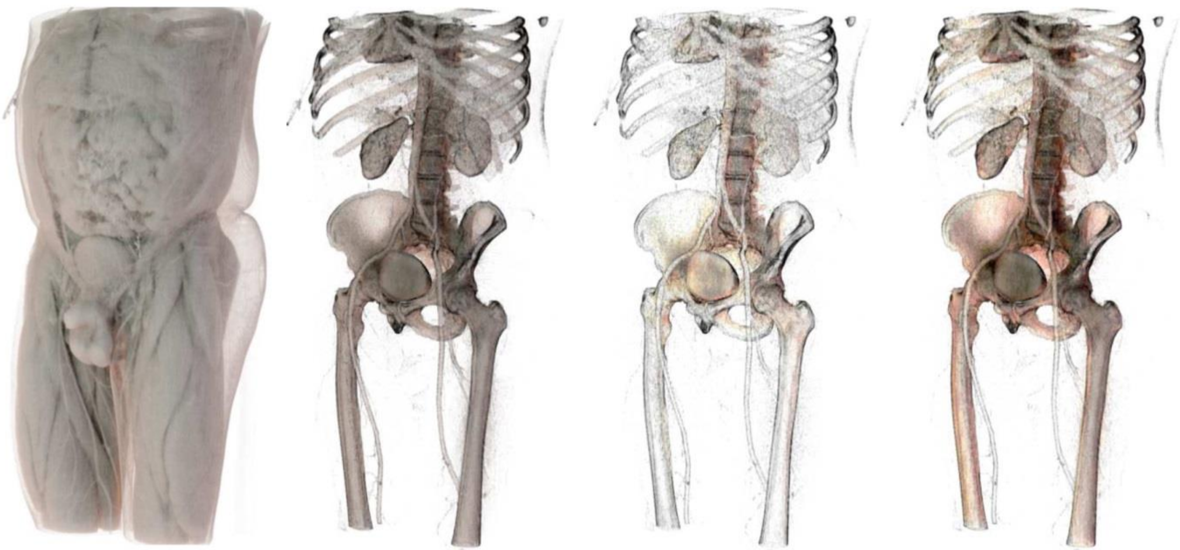
- Additive term to local lighting



Original | Vicinity shading, Stewart 2003 | **VOMI**

Color ambient occlusion





CAO map | CAO map with contours | CAO maps with contours and color quantization

Opacity modulation

Original

Modulated to emphasize skeleton

Original

Modulated to
emphasize ribs

Viewpoint selection

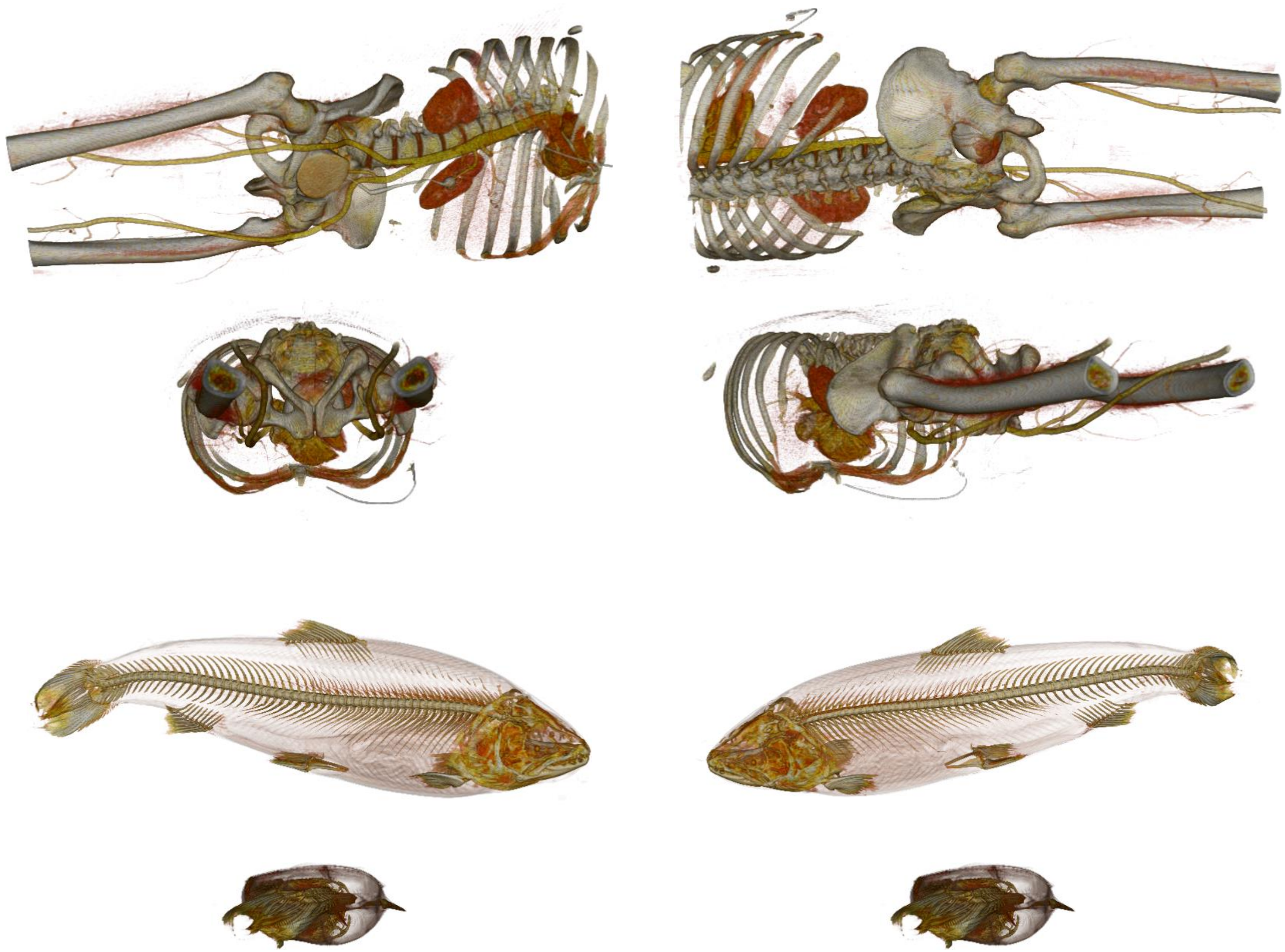
- VMI versus Informativeness

Min VMI | Max INF

Max VMI | Min INF

Min VMI | Max INF

Max VMI | Min INF



References

- T.M. Cover and J.A. Thomas. Elements of Information Theory. Wiley, 1991, 2006
- R.W. Yeung. Information Theory and Network. Springer, 2008
- M.R. DeWeese and M. Meister. How to measure the information gained from one symbo., Network: Computation in Neural Systems, 10, 4, 325-340, 1999
- D.A. Butts. How much information is associated with a particular stimulus?. Network: Computation in Neural Systems, 14, 177-187, 2003
- J. Burbea and C.R. Ra. On the convexity of some divergence measures based on entropy functions. IEEE Transactions on Information Theory, 28, 3, 489-495, 1982
- Noam Slonim and Naftali Tishby. Agglomerative Information Bottleneck. NIPS, 617-623, 1999

References

- Imre Csiszár and Paul C. Shields. Information Theory and Statistics: A Tutorial. Communications and Information Theory, 1, 4, 2004
- Pere P. Vazquez, Miquel Feixas, Mateu Sbert, and Wolfgang Heidrich. Viewpoint selection using viewpoint entropy. In Proceedings of Vision, Modeling, and Visualization 2001 , pages 273-280, Stuttgart, Germany, November 2001.
- M. Sbert, M. Feixas, J. Rigau, M. Chover, I. Viola. Information Theory Tools for Computer Graphics. Morgan and Claypool Publishers, 2009
- Bordoloi, U.D. and Shen, H.-W. (2005). View selection for volume rendering. In IEEE Visualization 2005 , pages 487-494
- Ji, G. and Shen, H.-W. (2006). Dynamic view selection for time-varying volumes. Transactions on Visualization and Computer Graphics , 12(5):1109-1116

References

- Mühler, K., Neugebauer, M., Tietjen, C. and Preim, B. (2007). Viewpoint selection for intervention planning. In Proceedingss of Eurographics/ IEEE-VGTC Symposium on Visualization, 267-274
- Ruiz, M., Boada, I., Feixas, M., Sbert, M. (2010). Viewpoint information channel for illustrative volume rendering. Computers & Graphics , 34(4):351-360
- Takahashi, S., Fujishiro, I., Takeshima, Y., Nishita, T. (2005). A feature driven approach to locating optimal viewpoints for volume visualization. In IEEE Visualization 2005 , 495-502
- Viola, I, Feixas, M., Sbert, M. and Gröller, M.E. (2006). Importance-driven focus of attention. IEEE Transactions on Visualization and Computer Graphics , 12(5):933-940

Thanks for your attention!