Tutorial

Bidirectional Texture Functions Modelling

Visual Quality Verification

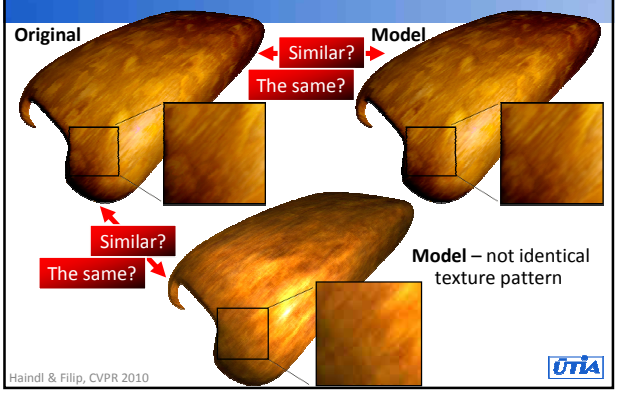
Michal Haindl Jiří Filip

Institute of Information Theory and Automation
of the AS CR

13/06/2010
CVPR 2010

www.utia.cz

Motivation of Visual Verification



Principles of Visual Verification

- Methods producing pixel-wise reproduction of original data ⇨ standard metrics used in image restoration (L1, L2, PSNR, SSIM, VDP).
- Statistical methods produce only similar texture ⇨ measures defining visual similarity between two different textures.
  - ~~Attempts based on second [Julezs 1962] and third order statistics [Yellott 1993] (not metric).~~
  - ~~Approaches based on combinations of cumulative histograms, Gabor features [Manjunath & Ma 96], steerable pyramids [Portilla & Simoncelli 00], Local Binary Patterns [Ojala et al. 02], Markov Random Fields features [Vacha & Haindl 07], [Filip et al. 10].~~
  - However, the most reliable way is visual psychophysics.

Principles of Visual Verification

- Comparing visualizations of original data vs. model
- Requirements:
  - Fast similarity evaluation
  - In agreement with human visual perception
  - Ability to evaluate similarity of different textures
- Visual Quality Assesment Metrics
  - Metrics L1, L2, PSNR, SSIM, VDP
- Visual Psychophysics

Image Quality Assessment Metrics

- Measure visual similarity of two textures
  - Original vs. modified image (model)
- Advantages
  - Automatic and fast results evaluation
  - Provide similarity of pixel-wisely registered images & difference maps
- Disadvantages
  - Pixel-wise comparison, e.g., not translation invariant
  - Do not provide general similarity of two different images

PSNR - Peak Signal-to-Noise Ratio

- an approximation to human perception of reconstruction quality in compression codecs
- Based on Mean Square Error (MSE):

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}||I(i,j) - K(i,j)||^2$$

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$

- The same images - PSNR = ∞ dB
- Acceptable range for video coding 30-50 dB

SSIM – Structure Similarity Index Metric

[Wang et al. 04]
- compares **luminance**, **contrast**, and **structure**
- local window (11x11pixels), Gaussian weighting
- range (-1,1), 1 – images are the same

$$S(\mathbf{x},\mathbf{y}) = f[l(\mathbf{x},\mathbf{y}), c(\mathbf{x},\mathbf{y}), s(\mathbf{x},\mathbf{y})]$$

$$l(\mathbf{x},\mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \qquad c(\mathbf{x},\mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \qquad s(\mathbf{x},\mathbf{y}) = \frac{2\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

$$S(\mathbf{x},\mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

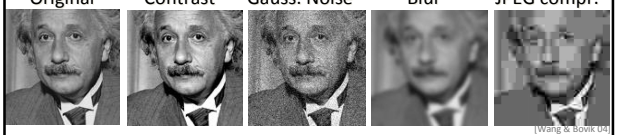
$$C_1 = (K_1 L)^2 \qquad C_2 = (K_2 L)^2 \qquad C_3 = C_2/2$$

$$K_1 = 0.01 \text{ and } K_2 = 0.03$$

SSIM – Structure Similarity Index Metric

Z. Wang, A. C. Bovik, Modern Image Quality Assessment

Original | Contrast | Gauss. Noise | Blur | JPEG compr.



[Wang & Bovik 04]

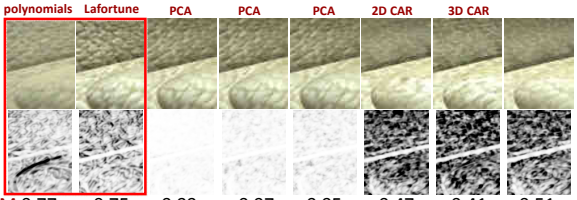
MSE	306	309	308	309
PSNR [dB]	23.3	23.2	23.2	23.2
SSIM	0.93	0.58	0.64	0.58
CW-SSIM	0.94	0.81	0.60	0.63

CW-SSIM – Complex Wavelet Domain SSIM
insensitive to: - luminance/contrast changes and small geometrical distortions such as translation, scaling, and rotation

SSIM – BTF Compression Methods

Original data

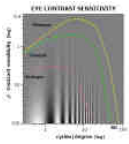
Per-view polynomials	Per-view Lafortune	Per-view PCA	Entire PCA	Per-cluster PCA	MRF 2D CAR	MRF 3D CAR	GMRF
SSIM 0.77	0.75	0.99	0.97	0.95	0.47	0.41	0.51

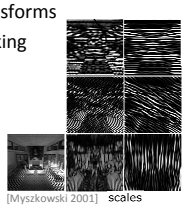


VDP – Visual Difference Predictor

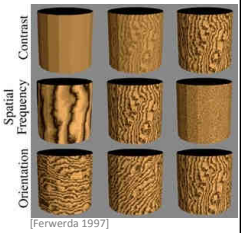
[Daly 93] -- Model of low-level human vision for a given viewing conditions (screen resolution & size, observer distance)

- Basic concepts:
  - Contrast Sensitivity Function
  - Cortex Transforms
  - Visual Masking





[Myszkowski 2001]

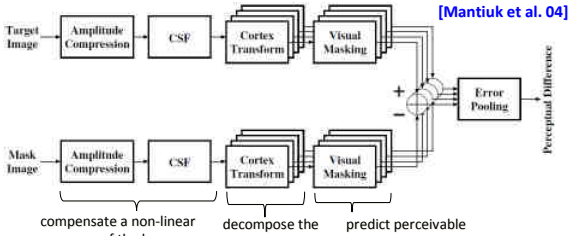


[Ferwerda 1997]

VDP – Visual Difference Predictor

- Target and mask image processing pipeline

[Mantiuk et al. 04]



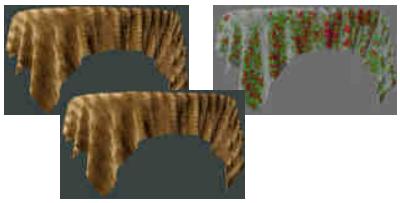
compensate a non-linear response of the human eye to luminance and the loss of sensitivity

decompose the image into spatial and orientational channels

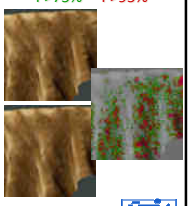
predict perceivable differences in each channel separately

VDP Applications

[Mantiuk et al. 04] – HDR implementation
http://www.mpi-inf.mpg.de/resources/hdr/vdp/

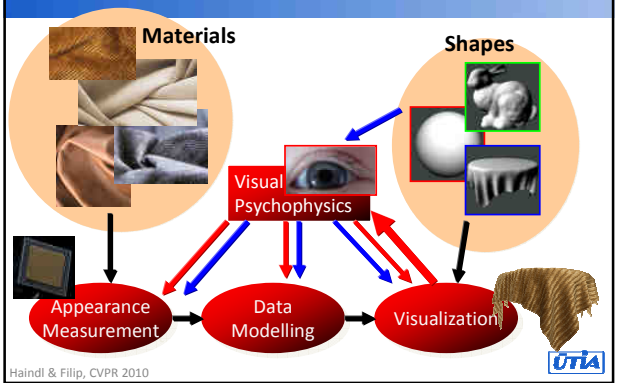


Pixel visual difference probability
P>75% P>95%



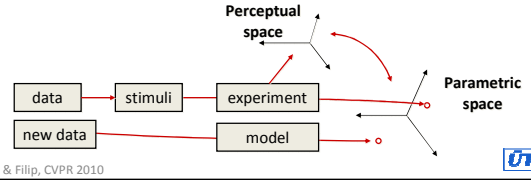
- High-quality rendering stopping
- Adaptive mesh subdivision
- Compression artifacts evaluation

Visual Psychophysics



Visual Psychophysics

- What we can get from it?
  - Perceptual validation
  - Visual comparison of different methods
  - Tune methods parameters in accordance with human vision
  - Help to find perceptually relevant computational model



Psychophysical Experiment Design

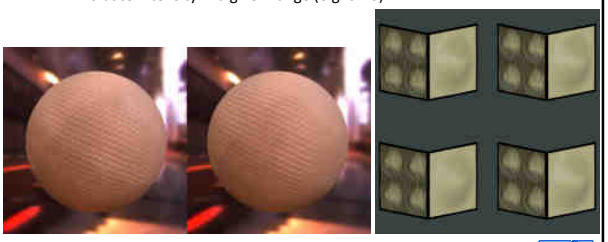
- Which visual feature or underlying parameter are analyzed? ⇨ Stimuli preparation.
- What result are expected? What type of data analysis we assume? ⇨ Data collection.
- Subjects: number, background, motivation (paid vs. volunteers), age? ⇨ Consistency.
- What will be the subjects' task?

  Task formulation is crucial. ⇨ Consistency.
- Conditions of the experiment and its length?

  Control as much factors as possible.

Experimental Stimuli - static

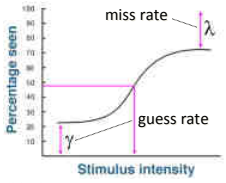
- Static stimuli - comparing of images
  - Yes/No answer
  - Pick up different object
  - Evaluate intensity in a given range (e.g. 0-10)



Psychophysical Data Analysis

- Subjects results **normalization and pooling** (average, median, etc.)
- Are the results **statistically significant**? ⇨ statistical tests (ANOVA, Kruscal-Wallis, Cochran, etc.)
- **Psychometric function** ⇨ dependence of subjects' response on stimulus intensity

[Wichmann & Hill, 2001]



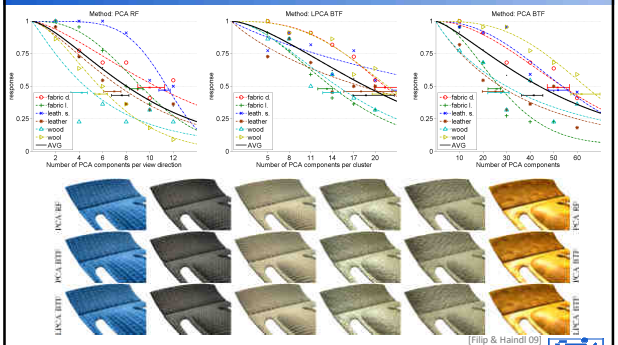
$$\psi(x;\alpha,\beta,\gamma,\lambda) = \gamma + (1-\gamma-\lambda)\,F(x;\alpha,\beta)$$

$$F(x;\alpha,\beta) = 1-\exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right), \quad x \geq 0$$

$$\psi_p^{-1} = \alpha\sqrt[\beta]{\ln\left(\frac{1-\gamma-\lambda}{1-p-\lambda}\right)}$$

p is the desired percentage (usually 50%)

Example: Methods Quality Evaluation

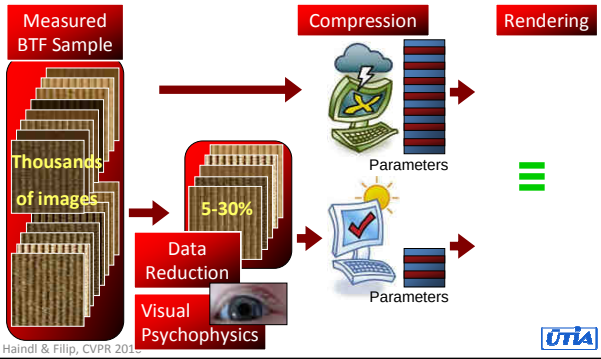


[Filip & Haindl 09]

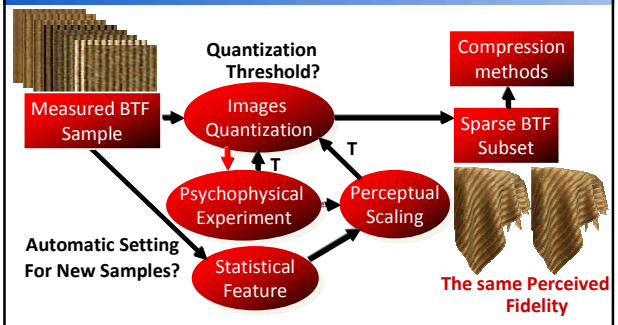
[Filip & Haindl, IEEE PAMI 2009]

Example: BTF Data Reduction

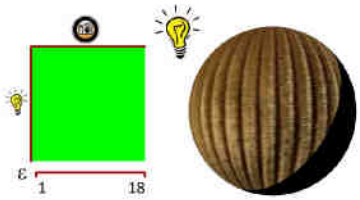
[Filip et al. ACM TOG 08]



Method Overview

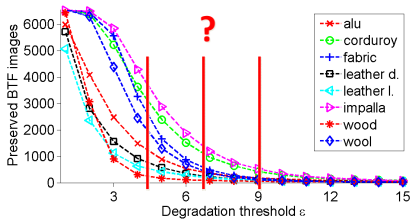


BTF Vector Quantization – demo



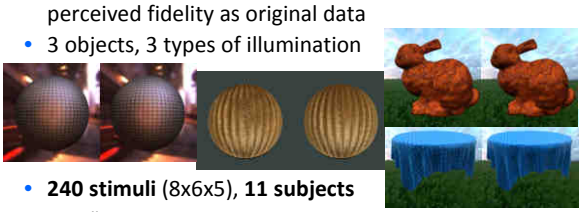
BTF Vector Quantization

- No. of **preserved images** vs. **quantization threshold**



Psychophysical Experiment

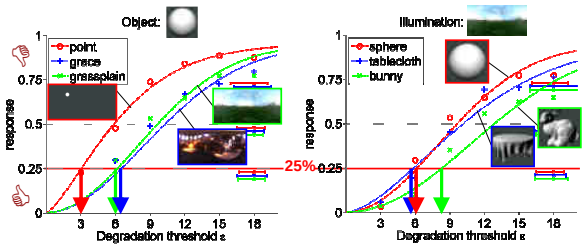
- **Goal:** Find quantization thresholds giving the same perceived fidelity as original data
- 3 objects, 3 types of illumination



- **240 stimuli** (8x6x5), **11 subjects**
- 20.1" LCD calibrated screen
- *'Can you detect any differences in the materials?'*

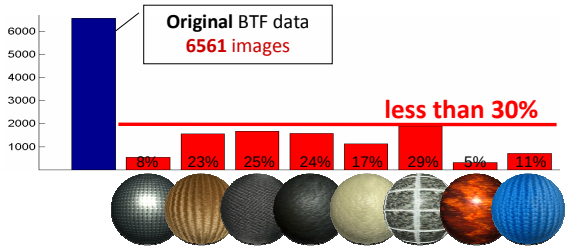
Psychophysical Experiment

- Different **illuminations**
- Different **objects**

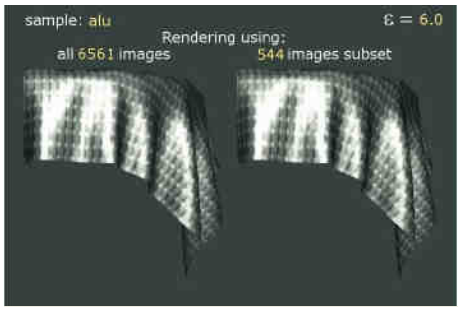


BTF Perceptual Subset

- **Numbers** of preserved BTF **images** for **estimated thresholds**

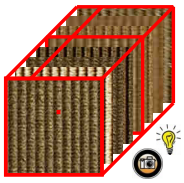


Results – Psychophysical Thresholds



Psychophysically-Based Metric

- **Automatic metric** for data reduction **needed**
- Data variance analysis – candidates for metric:
  - σ_1 **total variance** in BTF data
  - σ_2 mean **variance** across **illuminations and views**
  - σ_3 mean **variance** across **image** planes

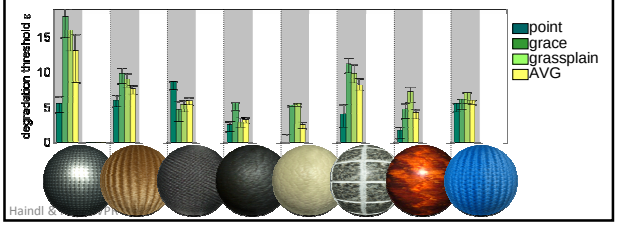


Psychophysically-Based Metric

Set of linear equations: $\begin{bmatrix} \hat{\varepsilon}^1 \\ \vdots \\ \hat{\varepsilon}^n \end{bmatrix} = s \begin{bmatrix} {\sigma_3}^1 \\ \vdots \\ {\sigma_3}^n \end{bmatrix}$

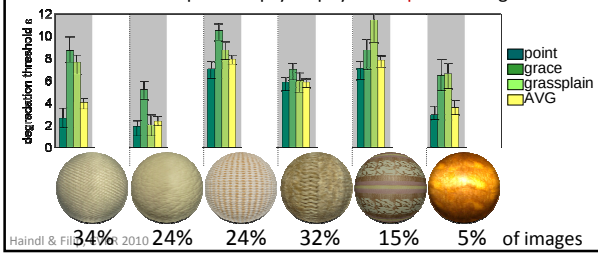
New threshold: $\varepsilon_{\sigma_3} = s\sigma_3$

- σ_3 (**R=0.76**) mean variance across image planes

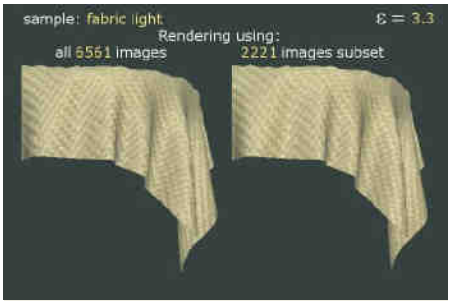


Validation Experiment

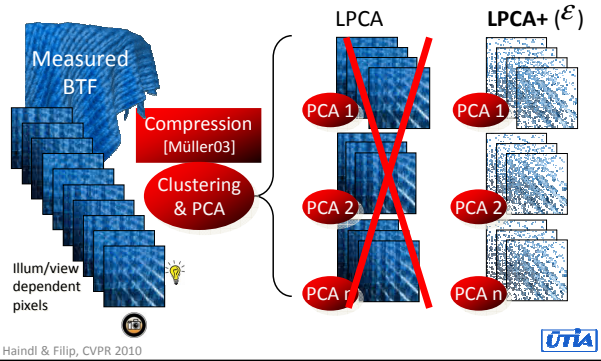
- **6 additional BTF samples, 18 subjects**, one shape
- scaling value s from the previous experiment.
  - Automatic thresholds setting metric
  - No need to perform psychophysical experiment again



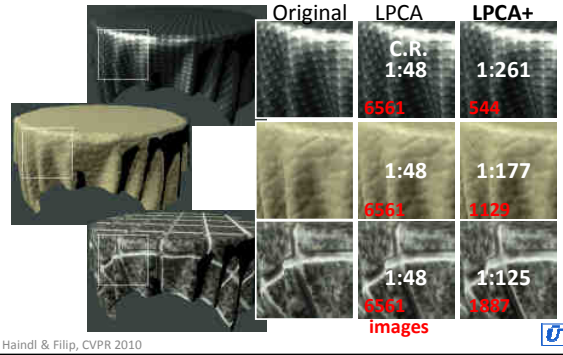
Results – Derived Thresholds



Applications – BTF Compression

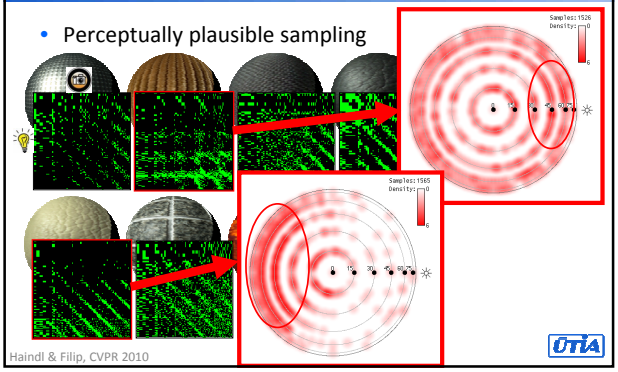


Applications – BTF Compression

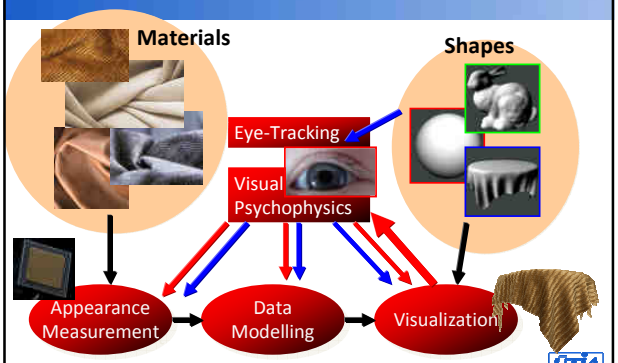


Applications – BTF Sampling

- Perceptually plausible sampling



Visual Psychophysics



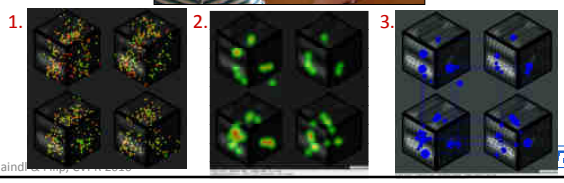
Gaze Data Recording (Eye-Tracking)

Task: *Which one is different?*

Gaze recording device (Tobii x50)
- Binocular
- Infrared-based

Stimulus image

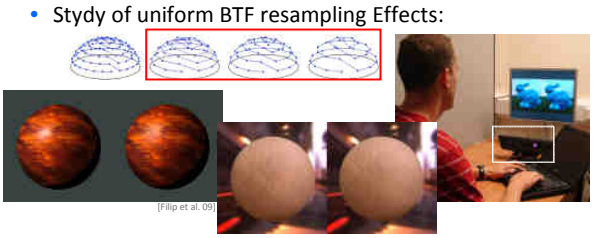
1. Gaze fixations locations a durations
2. Hot-spot maps
3. Visual search patterns



Study of Resampling Effects

[Filip et al. ACM TAP 2009]

- Stydy of uniform BTF resampling Effects:



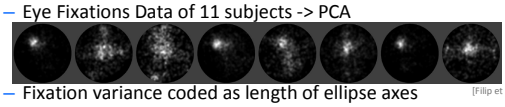
- Subjects responses ⇨ resampling is less apparent:
  - in azimuthal angles than in elevation ones
  - in illumination directions than in view directions
  - for environmental illumination & less glossy samples

Example: Gaze Analysis of Materials

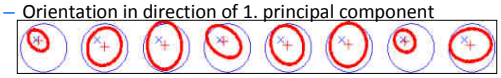
[Filip et al. ACM TAP 2009]

- Analysis of Gaze Data for Different Samples

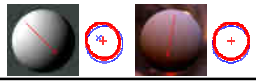
  - Eye Fixations Data of 11 subjects -> PCA



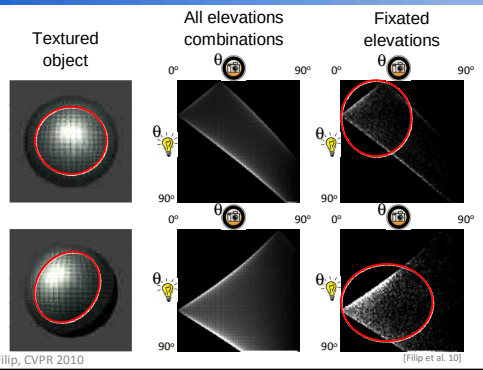
  - Fixation variance coded as length of ellipse axes [Filip et al. 09]
  - Orientation in direction of 1. principal component



- Slant of main axis = illumination gradient



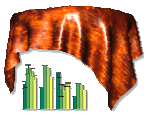
Example: Gaze Analysis of Materials



[Filip et al. 10]

Conclusions on Visual Verification

- Applicable metrics exist for pixel-wise comparison.
- Lack of reliable general texture similarity metric
- Visual Psychophysics allows:
  - Arbitrary customized stimuli.
  - Time demanding & expensive, but can provide valuable results.
  - Enables to find optimized parametric setting.
  - Useful for identification of proper statistical features ⇨ this can be difficult.
- Applications shown
  - More efficient compression and visualization
  - Selective material sampling
  - BTF visual attention prediction



References

Texture similarity

- B. Julesz. *Visual pattern discrimination*. IRE Transactions on Information Theory, pages 84–92, February 1962,
- I J. Yellot, John I. Implications of triple correlation uniqueness for texture statistics and the Julesz conjecture. Journal of the Optical Society of America A, 10(5):777–793, May 1993.

Structure similarity index (SSIM)

- Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, *Image quality assessment: From error visibility to structural similarity*," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004,
- Z. Wang, A. C. Bovik, *Modern Image Quality Assessment*, Morgan & Claypoll Publishers, 2006.

Visual difference predictor (VDP)

- Daly, S.: *The visible differences predictor: an algorithm for the assessment of image fidelity*., Digital Images and Human Vision, MIT Press, pp. 179--206, 1993.
- Mantiuk, R. and Myszkowski, K. and Seidel, H.-P.: *Visible Difference Predicator for High Dynamic Range Images*, IEEE International Conference on Systems, Man and Cybernetics, 2004

References

Steerable pyramids

- J. Portilla and E. P. Simoncelli. *A parametric texture model based on joint statistics of complex wavelet coefficients.* International Journal of Computer Vision, 40(1):49–71, 2000.

Local Binary Patterns

- Ojala, T., Pietikainen, M., Maenpaa, T.: *Multiresolution gray-scale and rotation invariant texture classication with local binary patterns*. IEEE PAMI 24(7) (2002) 971-987.

Gabor features

- B. S. Manjunath and W. Y. Ma. *Texture features for browsing and retrieval of image data.* IEEE PAMI, 18(8):837–842, August 1996,
- Bovik, A.: *Analysis of multichannel narrow-band lters for image texture segmentation*. IEEE Trans. on Signal Processing 39(9) (1991) 2025-2043.

Markov Random Field features

- Vacha, P., Haindl, M.: *Image retrieval measures based on illumination invariant textural MRF features*. In: CIVR, ACM (2007) 448-454,
- Haindl, M., Simberova, S.: *A Multispectral Image Line Reconstruction Method. In: Theory & Applications of Image Analysis*. World Scientic Publishing Co., Singapore (1992) 306-315.

References

Visual psychophysics

- Filip J., Haindl M.: *Bidirectional Texture Function Modeling: A State of the Art Survey*. IEEE TPAMI, 31(11), November 2009, pp. 1921-1940,
- Filip J., Chantler M.J., Green P.R., Haindl M.: *A Psychophysically Validated Metric for Bidirectional Texture Data Reduction*. ACM Transactions on Graphics 27(5) (proceedings of SIGGRAPH Asia 2008), Article 138, 11 pp.

Visual psychophysics + Eye-tracking

- Filip J., Chantler J. M., Haindl M.: *On Optimal Resampling of View and Illumination Dependent Textures*. ACM Transactions on Applied Perception, 6(3), August 2009, pp. 18,
- Filip J., Haindl. M, Chantler M.J.: *Gaze-Motivated Compression of Illumination and View Dependent Textures*. ICPR, August 2010,
- Filip J., Vacha P., Haindl. M, Green P.R.: *A Psychophysical Evaluation of Texture Degradation Descriptors*. S+SSPR, August 2010.