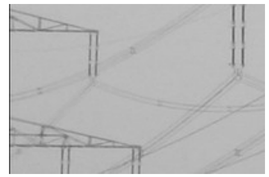
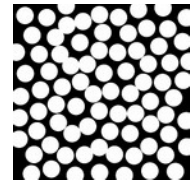
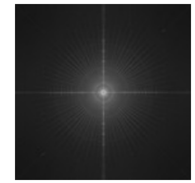
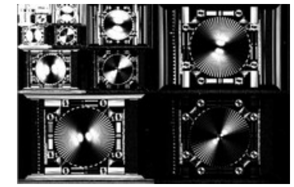
     

A short trip through Image Quality Analysis

Charles A. Wüthrich

CogVis/MMC -Faculty of Media

Bauhaus-Universität Weimar
Bauhausstr. 11

D-99423 Weimar

Germany

caw AT medien.uni-weimar.de

Overview

- Why image quality analysis? Motivation
- Available methods
- A first timid approach to using functional analysis for quality
  - FFT
  - Wavelets
- Analog vs. digital
- Which sampling grid makes sense?
- Conclusions and future work

How this all began…

- Take a high resolution scanner…

How this all began…

- Take a high resolution scanner…
- …forget a self printed analog photograph on the flatbed…

How this all began…

- Take a high resolution scanner…
- …forget a self printed analogue photograph on the flatbed…
- …think you put on the scanner a document to scan for income tax…

How this all began…

- Take a high resolution scanner…
- …forget a self printed analogue photograph on the flatbed…
- …think you put on the scanner a document to scan for income tax…
- …forget to choose reasonable settings for scanning….

How this all began…

- Take a high resolution scanner…
- …forget a self printed analogue photograph on the flatbed…
- …think you put on the scanner a document to scan for income tax…
- …forget to choose reasonable settings for scanning….
- …push the scan button…

Et Voila!

Tetenal Work paper Grade 3, scan @ 3600 dpi, 916.3 MB

Et Voila!

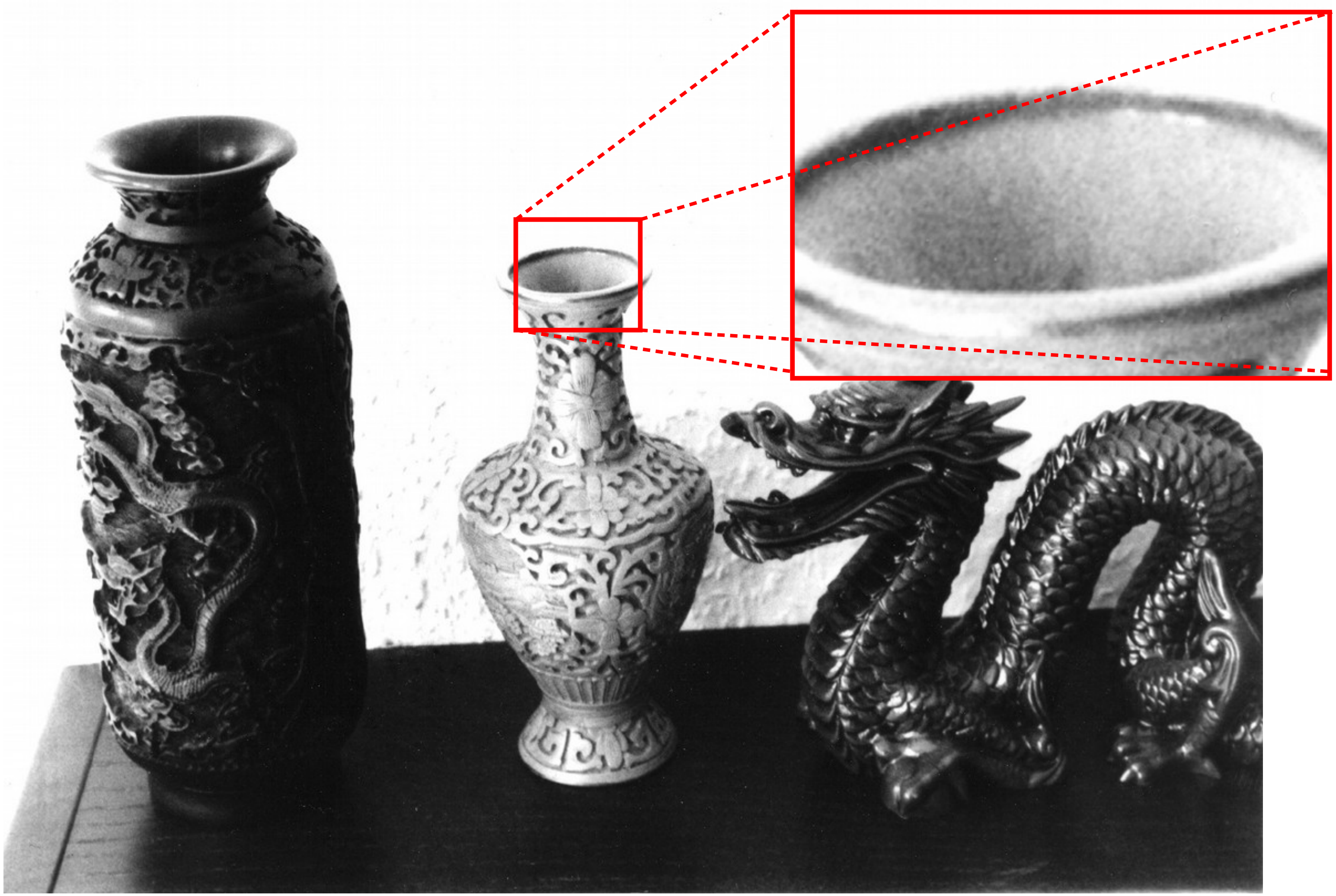
Definitely not a usable document for tax return!

Tetenal Work print paper grade 3, scan @ 3600 dpi, 916.3 MB

One Moment!

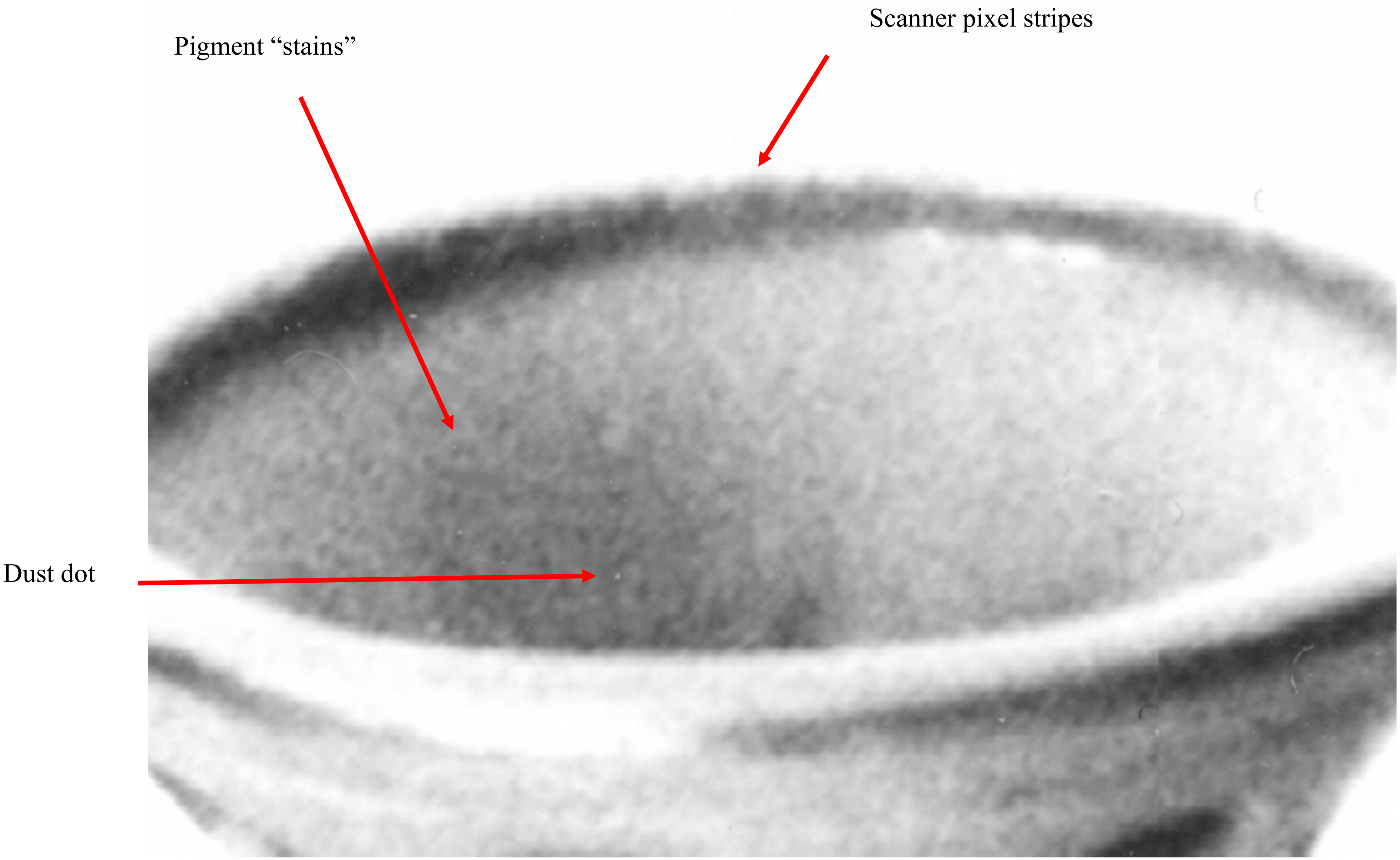
Tetenal Work print paper Grade 3, scan @ 3600 dpi, 916.3 MB!!!

One Moment!

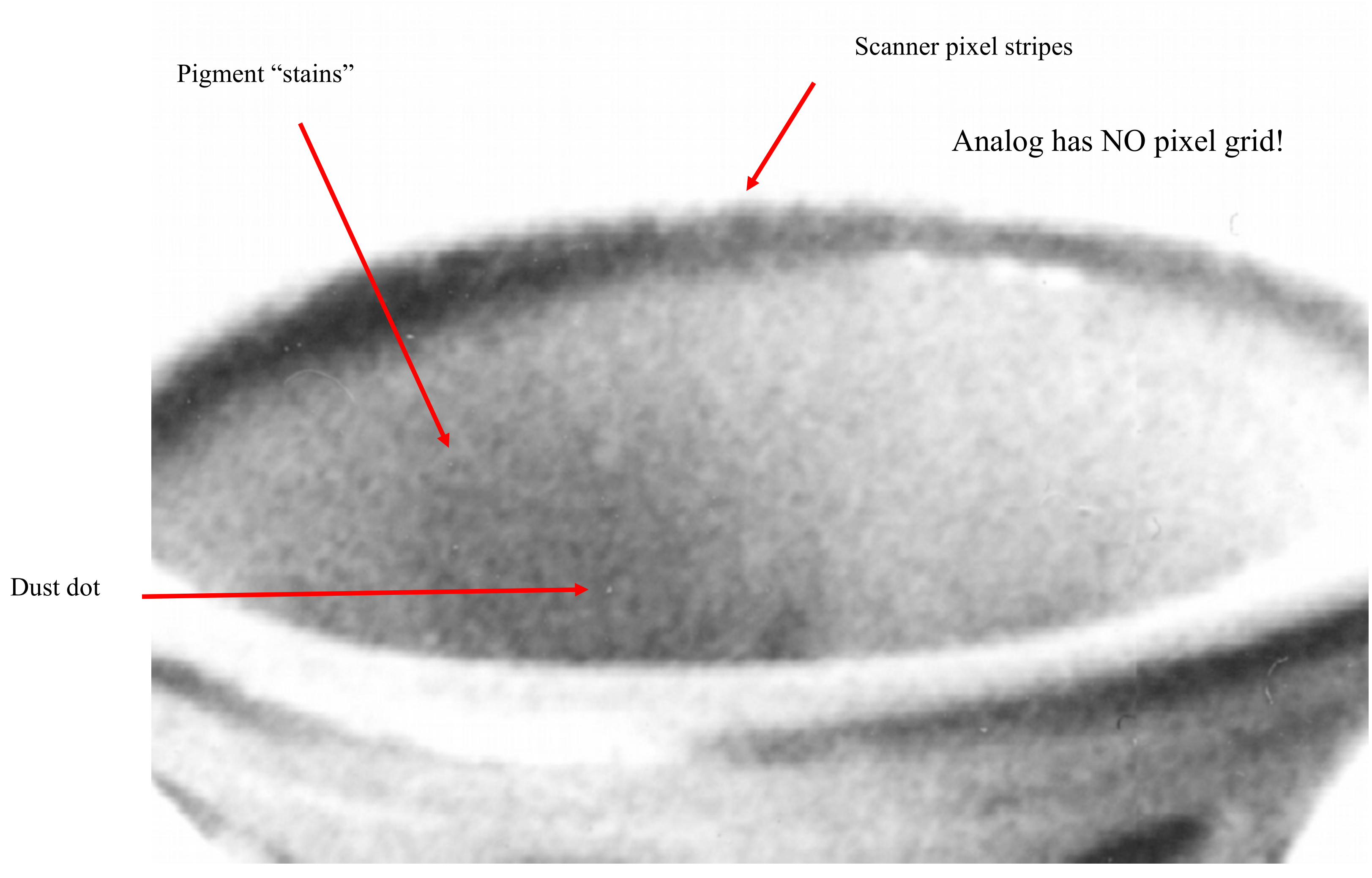


Tetenal Work paper Grade 3, scan @ 3600 dpi, 916.3 MB!!!

Enlargement



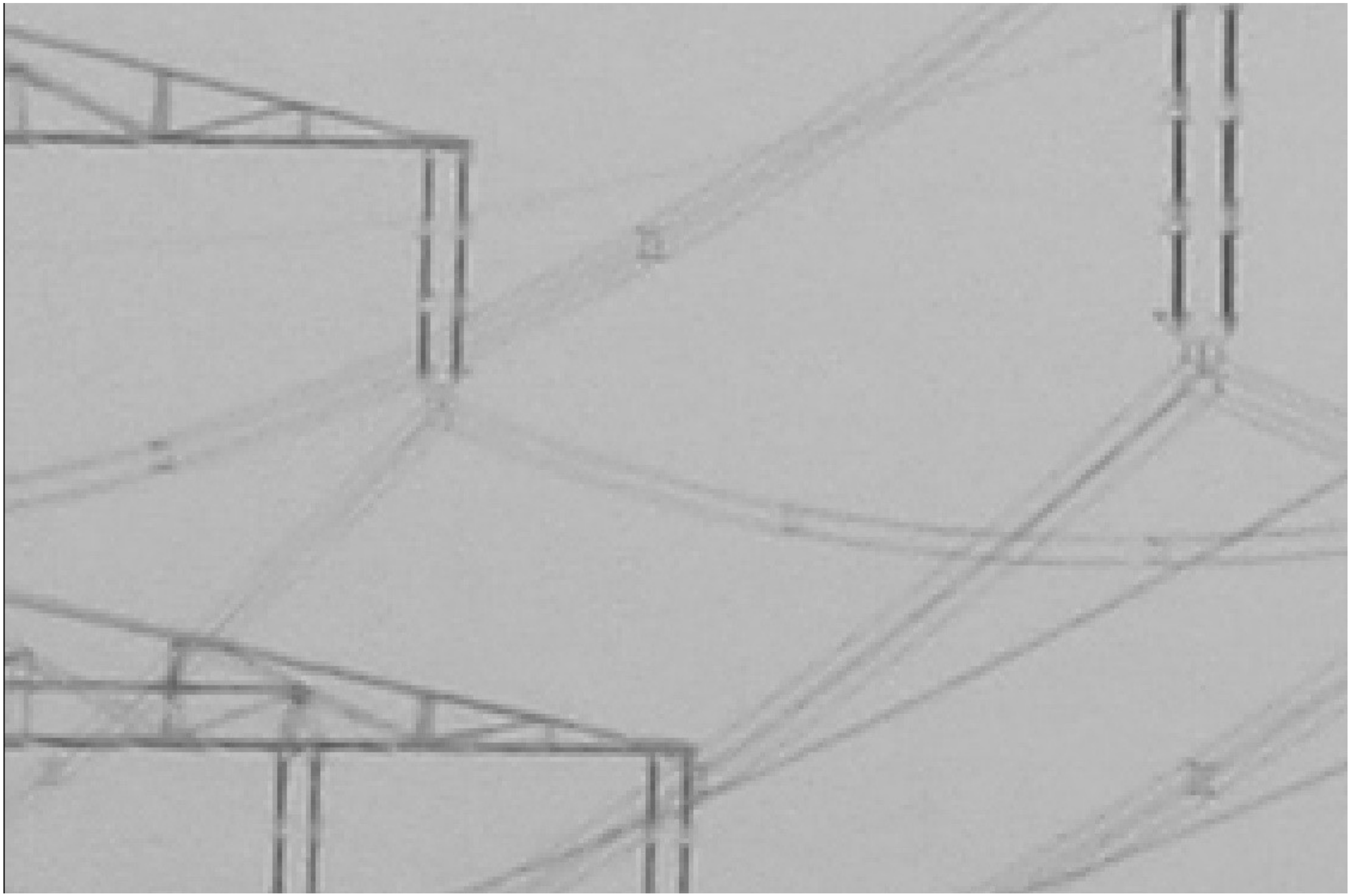
Print enlargement



What about analog film scans?

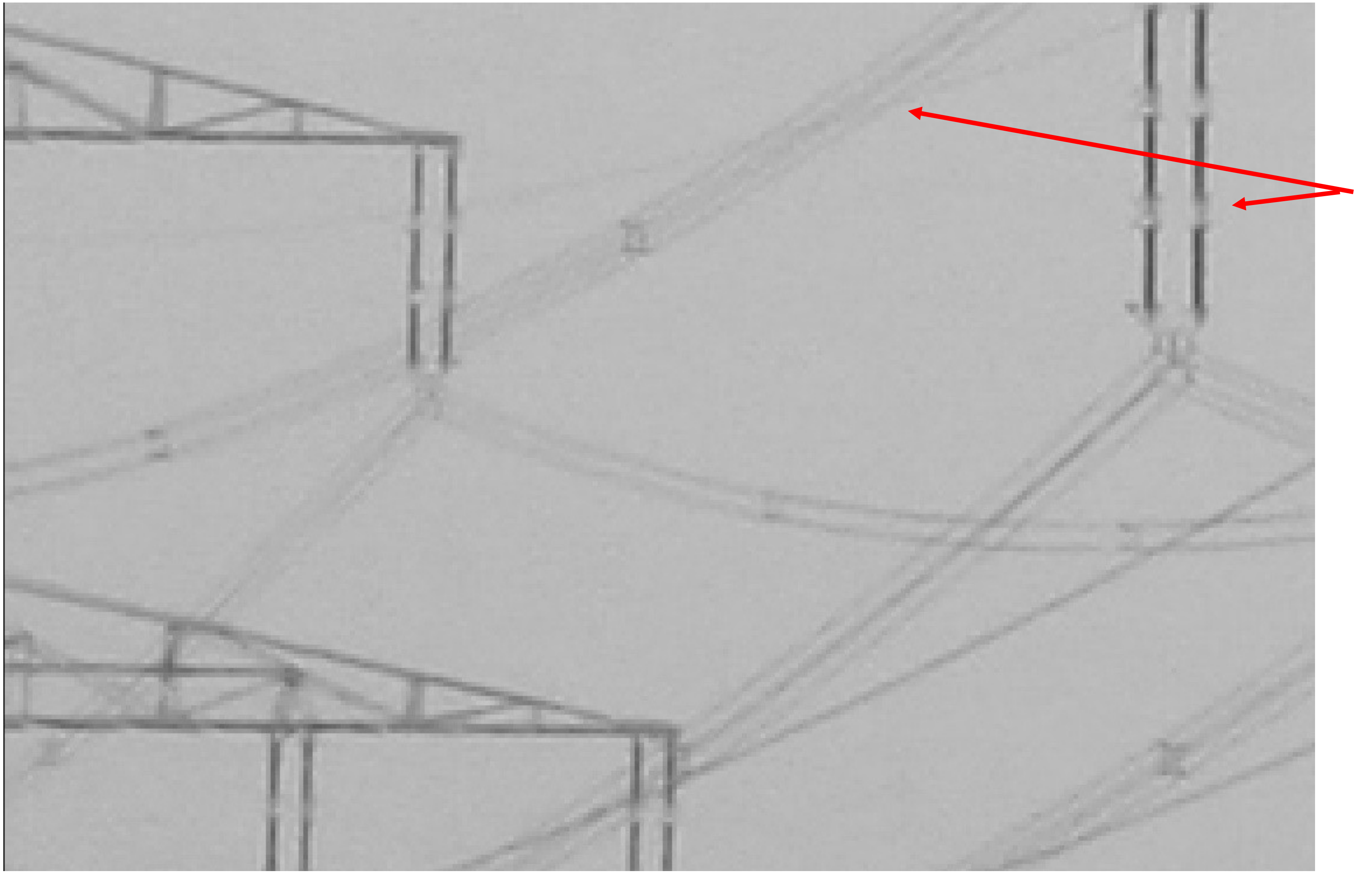
Fomapan 400, Rodinal, scan @ 1800 dpi

And digital pictures?



4x Enlargement, Canon 5D Mark II, ISO 400

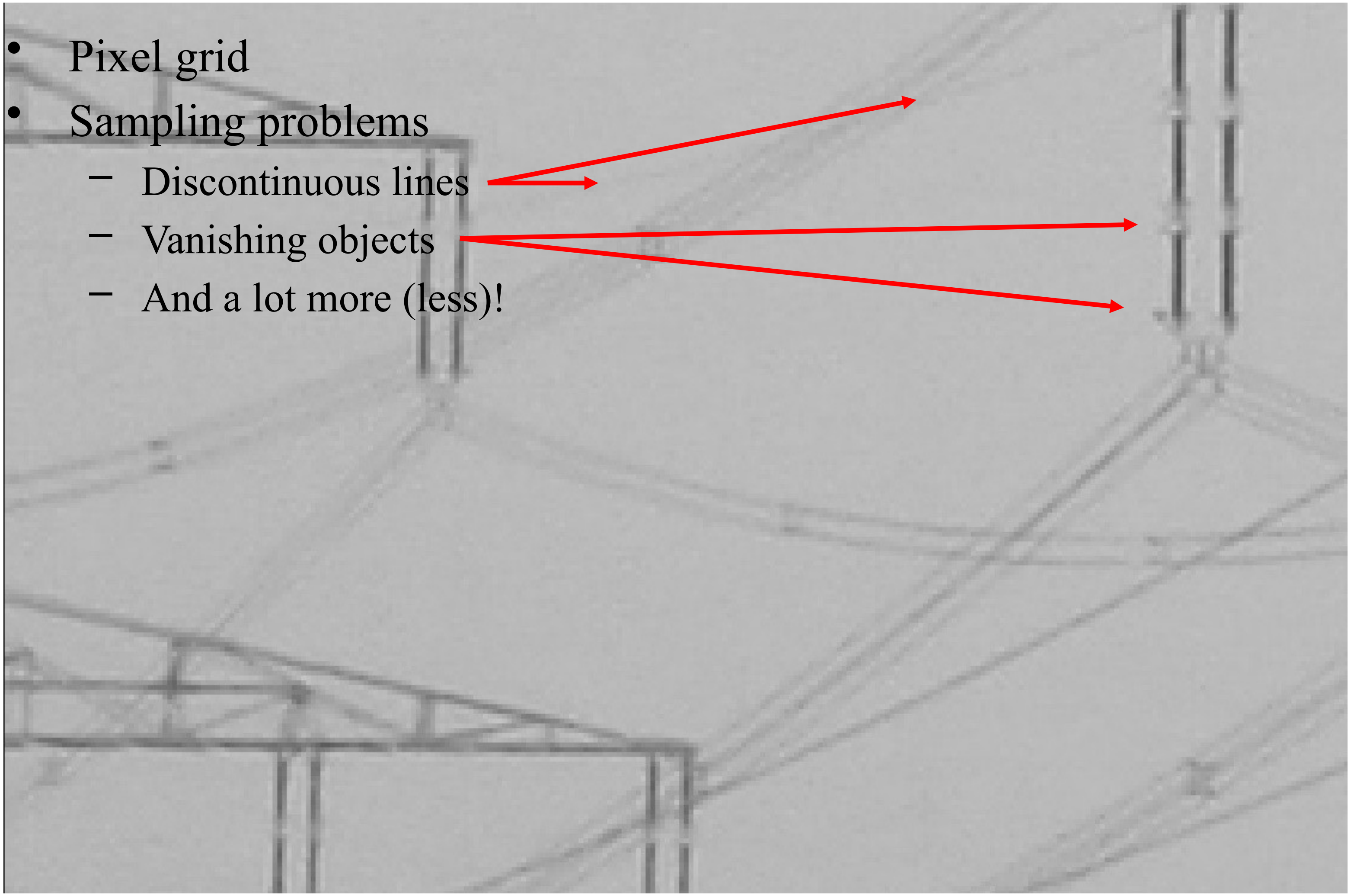
And digital pictures?



4x Enlargement, Canon 5D Mark II, ISO 400

And digital pictures?

- Pixel grid
- Sampling problems
  - Discontinuous lines
  - Vanishing objects
  - And a lot more (less)!



4x Enlargement, Canon 5D Mark II, ISO 400

Quality evaluation

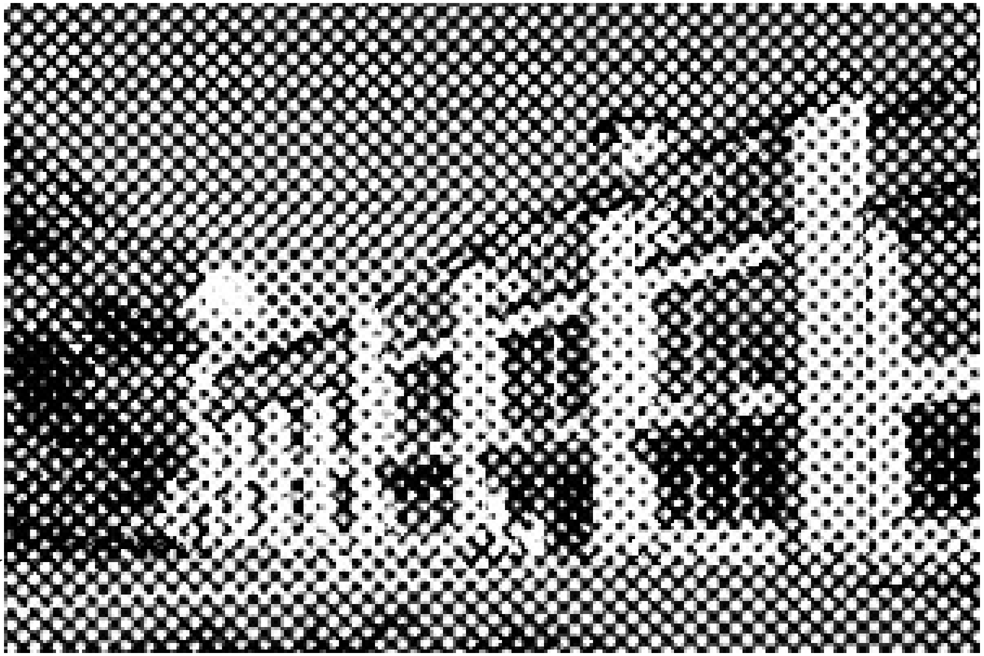
- Perhaps this is why artists say analog has still the edge?
  - Unfortunately, artists use often intuition and experience, but rarely can explain why it is like this.
- Quality assessment methods in the literature?
- Mostly geared towards
  - Image compression
  - Video compression
  - Thus enabling faster transmission, and lower storage space
- Defining quality not main purpose, but finding ways for evaluating it, eventually in real time
  - Methods are therefore simple

Quality evaluation

- To do evaluation:
  - Need samples
  - Need methods for evaluating
  - Combine methods into experiments
- The examples we will use:
  - Analog vs. digital
  - Different sampling grids and their performance at sampling: square, hexagonal, jittered.
  - Does graininess behave like noise?

Quality evaluation

- Image quality can be treated at different levels:
  - Microstructure level: Pixels and sampling
  - Higher level: Subject, Composition, Light, Colour
- But what we see is always seen by our visual system: perception plays a big role



Quality evaluation

- We will start from the microstructure level:
  Pixels and their dispositions
- Until now, the Imaging community used pixels placed at equal spacings on an orthogonal grid
- From the pictures of film photography we have learned that film did not have a grid for sampling light

- Question: Are there alternative pixel constellations that allow better sampling of light?
- How we define better quality?

Quality evaluation

- Image quality assessment methods:
  - Full reference (FR): reference image exists
  - No reference (NR): reference image does not exist
  - Reduced reference (RR): certain features are extracted from the reference image and employed by the quality assessment system as side information to help evaluate the quality of the distorted image

Sample generation

- Due to constraints, we used:
  - Digital and analog photographs comparison
    - Same subject, same optics, same aperture, same sensor size: 35mm film vs FF sensor
    - Digital photo vs Analog film (varying films and developer chemistry)
    - Scanned analog material at similar resolutions to FF camera
  - Sampling grid positions comparison using resampled digita images for alternative pixel dispositions:
    - Square sampling grid
    - Hexagonal sampling grid
    - Jitteed sampling
    - Randomized sampling (Voronoi partitions of the plane)

Quality evaluation: global methods

- Simple global measures used in Image Processing (simple):
  - Mean squared error (MSE)

$$MSE = \frac{1}{N}\sum_{i=1}^{n}(x_i - y_i)^2$$

  - Root of the mean square (RMS)
    - Incorporate variance and bias
    - Fails in many occasions, since it averages over whole picture
  - Signal and Peak Signal to Noise Ratio (SNR-PSNR) (Teo et Al.'94)

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

    where L=corrupting noise:
    - Addresses signal fidelity
    - Averages again
  - Shannon Entropy (Snannon'48, Tsai et Al.'08)

$$H = -\sum_{i=0}^{255} Pr(i) \cdot log_2 Pr(i)$$

- Commonly used in image transmission/compression

Quality evaluation from functional analysis

- Other methods use functional analysis and try to evaluate on the transformed space
  - Discrete Fourier Transform (DFT)
  - Discrete Cosinus Transform (DCT)
  - Discrete Wavelet Transform (DWT)
  - Complex Wavelet transform (CWT)
- In this lesson, we will use the DFT and CWT and test their usability to test image features with respect to quality
- To make things simpler, we will use black and white images

Quality evaluation from functional analysis

- In functional analysis, functions are approximated through a linear combination of basis functions $\beta_i(x)$:
  - They serve as functional coordinates to approximate a function *f(x)*
- Therefore, a function will be approximated as

$$f(x)=\Sigma_{-\infty,\dots,+\infty}\, c_i\, \beta_i(x)$$

- The Basis functions need to satisfy certain properties:
  - They must be elements of an inner product vector space V, in which an inner product

    $\langle,\rangle : V x V \rightarrow F$

    where F is a field, and for all vectors x,y,z ∈ V and scalars a ∈ F
    - $\langle x,y\rangle=\overline{\langle y,x\rangle}$ (conj. symmetry)
    - $\langle ax,y\rangle=a\langle x,y\rangle$ $\langle x+y,z\rangle=\langle x,y\rangle+\langle x,z\rangle$
    - $\langle x,x\rangle=0$ IFF x=0
  - The norm of a function x is defined as $\langle x,x\rangle$
  - They must be defined over the region of interest, and be orthonormal:
    - The inner product $\langle \beta_i(x),\beta_j(x)\rangle=0$ for i≠j.
    - The norm of each $\beta_i(x)$ is $\beta_i(x)=1\ \ \forall i$
- If these conditions are satisfied, then the basis functions can be used as “functional coordinates” for the the functional space

Functional analysis

- Fourier transforms, cosinus transforms, wavelet transforms and complex wavelet transforms differ only on the choice of the orthonormal basis:
  - Fourier transform: $\beta_i(\theta) = e^{2\pi i\theta} = \cos(2\pi\theta) + i \sin(2\pi\theta)$
  - Cosine transform: $\beta_i(\theta) = \alpha_p\alpha_q \cos(\pi(2m+1)p)/2M \cos(\pi(2n+1)q)/2N$, where $0 \leq p \leq M-1$ and $0 \leq q \leq N-1$
  - Wavelet transforms: the basis functions are Haar basis functions, i.e.
    - stretched and shifted versions of a bandpass function $\psi(t)$
    - Combined with shifts of a low pass scaling function $\phi(t)$
  - Complex wavelet transforms: here we have Haar complex basis functions

The tools: FFT analysis

- FFT analysis is suited at outlining
  - Differences in the frequency domain between images
  - Directional artifacts
- For the laymen, given *N* equidistant samples $x_0,\ldots,x_{N-1}$, their Discrete Fourier Transform is given by the N complex numbers

$$X_k = \sum_{n=0}^{N-1} x_n e^{\frac{2\pi i}{N} kn} \quad (k = 0, \ldots, N-1)$$

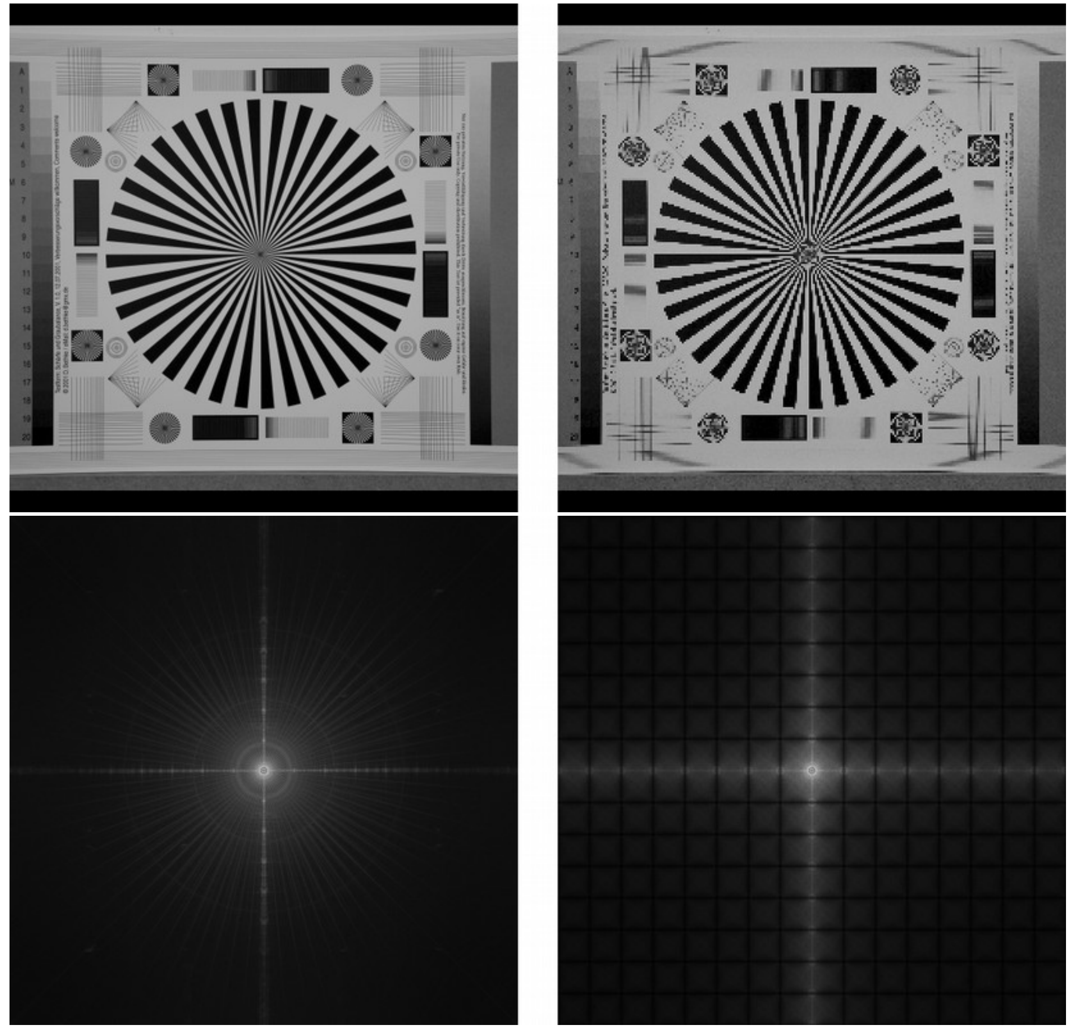
- Since an image is a bi-dimensional sample vector, the DFT of an image is obtained by performing its DFT to the rows, and then to the columns of the resulting bidimensional complex vector
- Usually, Cooley-Tuckey's algorithm FFT is applied, which works in logarithmic time, but requires images whose dimensions are a power of 2.

The tools: FFT analysis

- Once the DFT is applied, one can look at the
  - Amplitude: presence of a given frequency
  - Phase: offset with respect to edges
- We will take a look at amplitudes, representing the frequencies
  - Lower at center
  - High at corners
- The FFT treats images as periodic: shift invariant!
- Repetitive patterns
  - Aliasing

original

resampled: 16x16

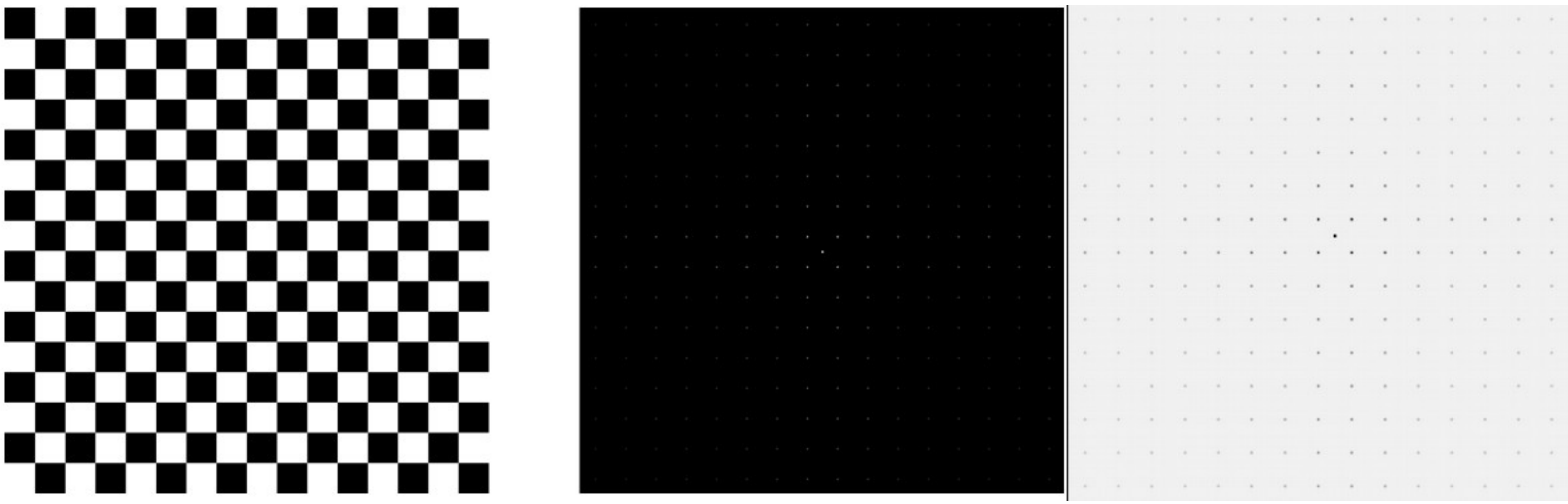


DFT, amplitude

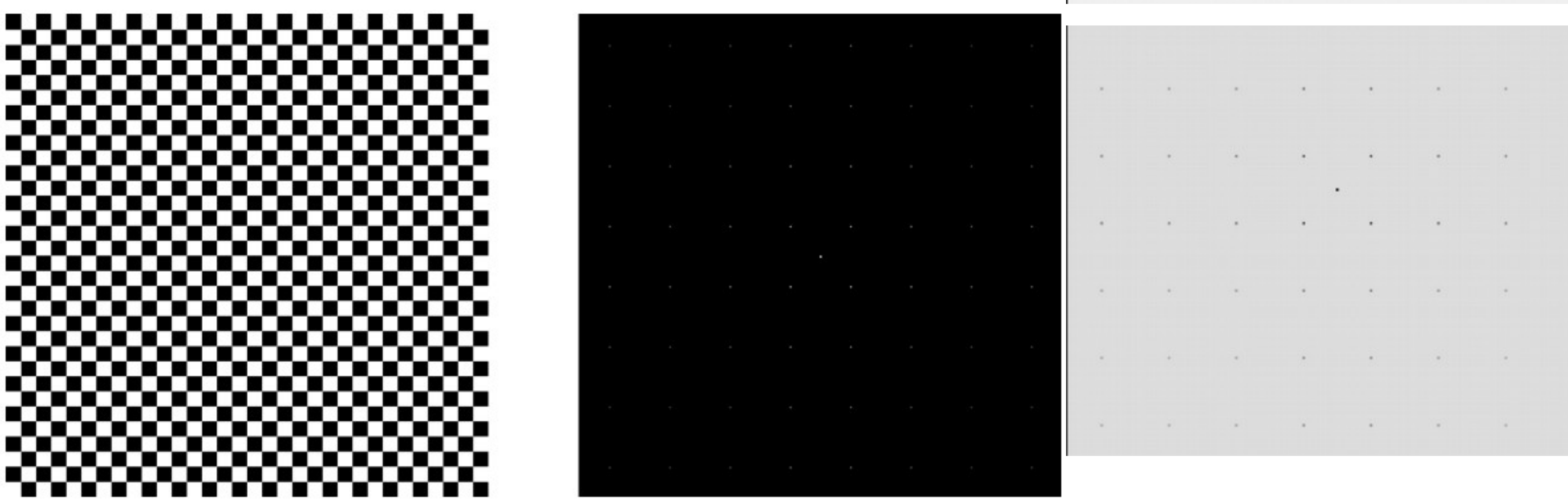
FFT amplitude characteristics

- Checkerboards show how the 2D FFT responds to frequencies:

8 cycles

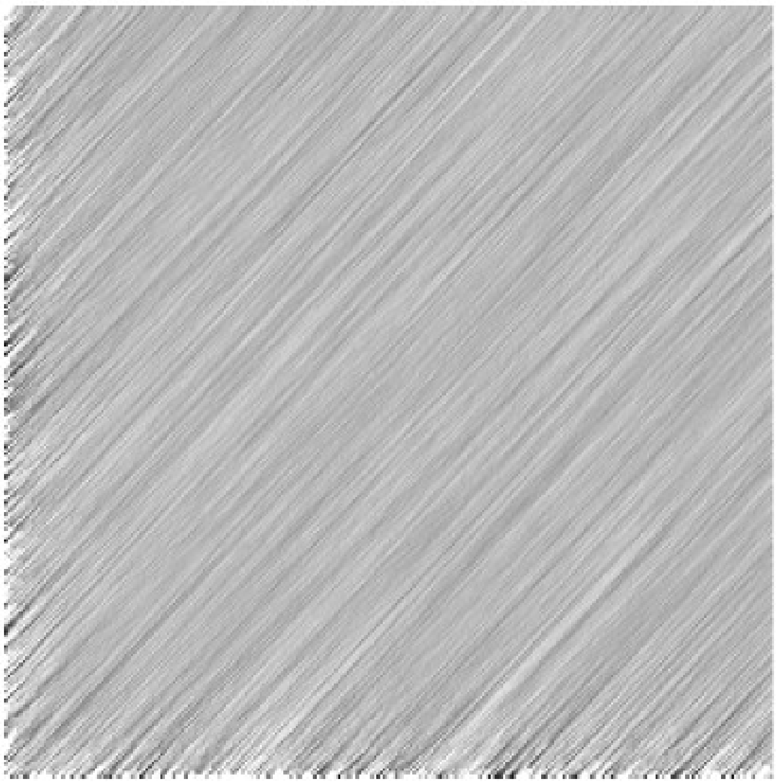


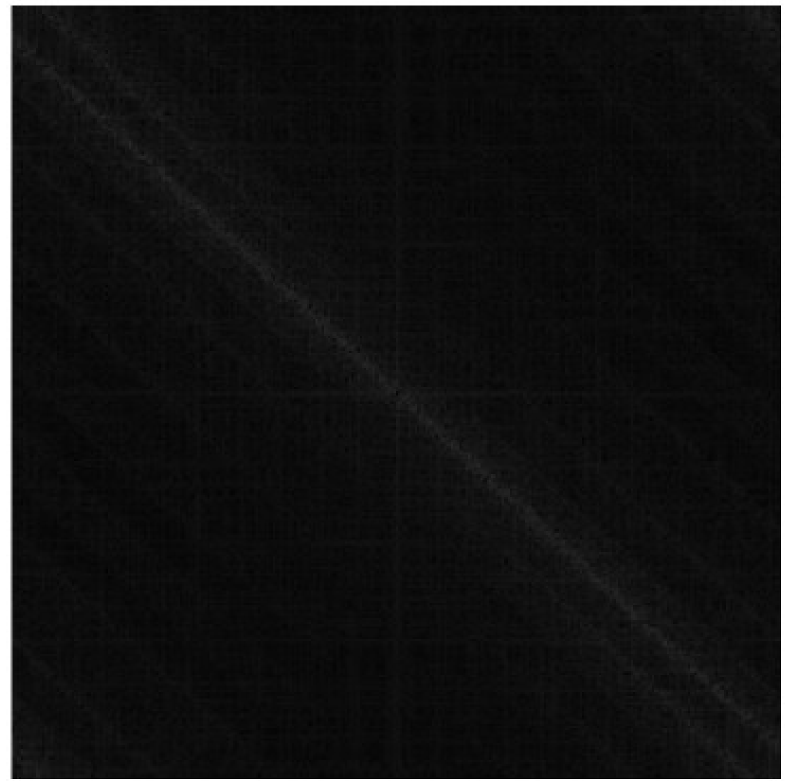
16 cycles



FFT amplitude characteristics

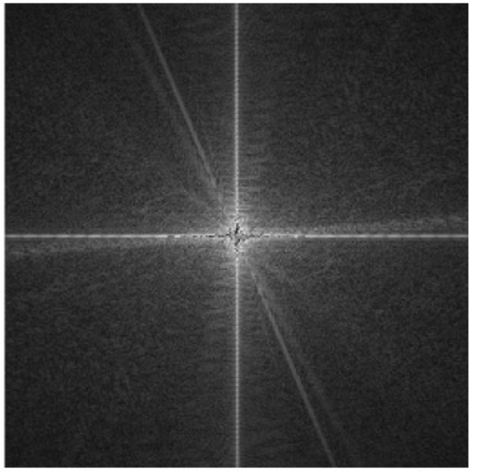
- Slanted lines show in the FFT a “line” perpendicular to the original slope

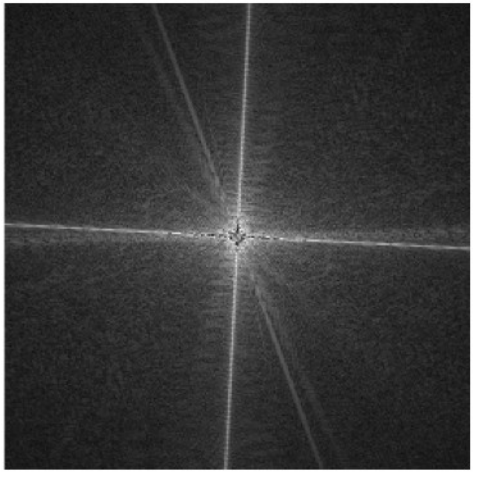




FFT amplitude characteristics

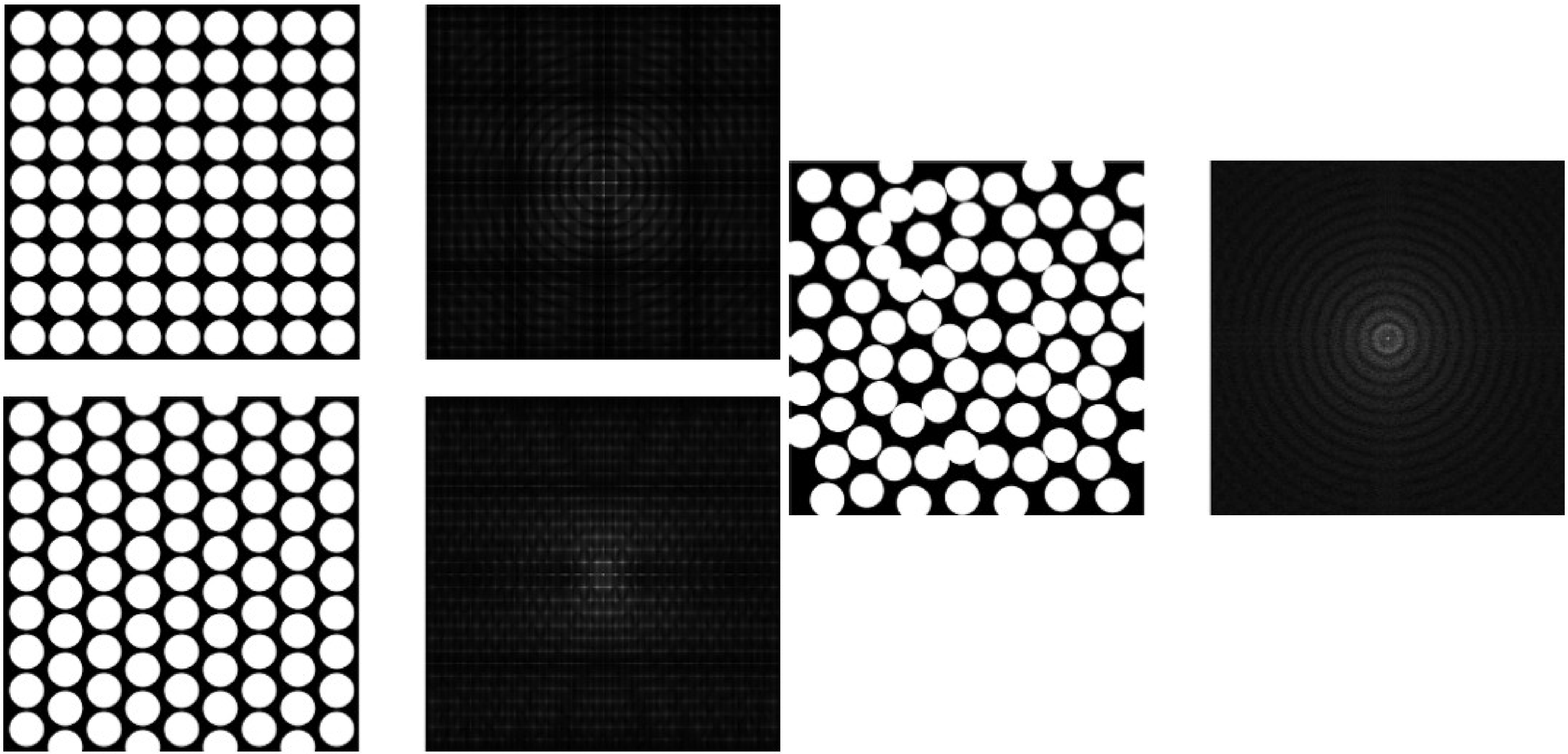
- Rotation of an image can be seen in the FFT quite clearly
- Second image is first image rotated
- Notice the slanting of the cross like pattern in the FFT





FFT amplitude characteristics

- Through the FFT one can evaluate sampling patterns and study their characteristics



Aperture and FFT

(a) f4.5

(b) f7.1

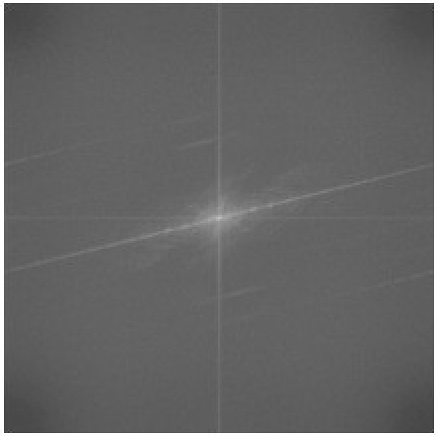
(c) f11

(d) f18

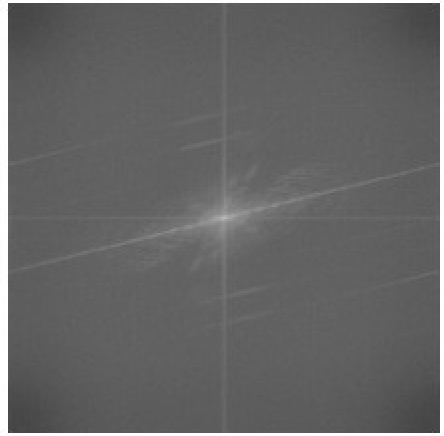
(e) f29

(f) f40

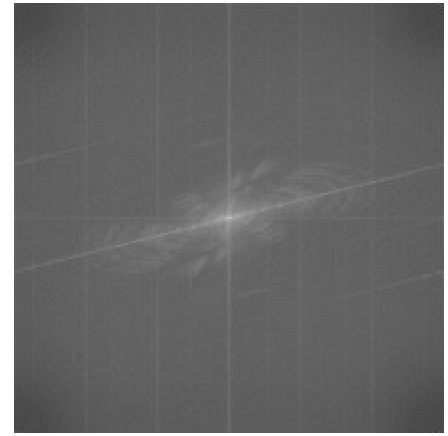
Aperture and FFT



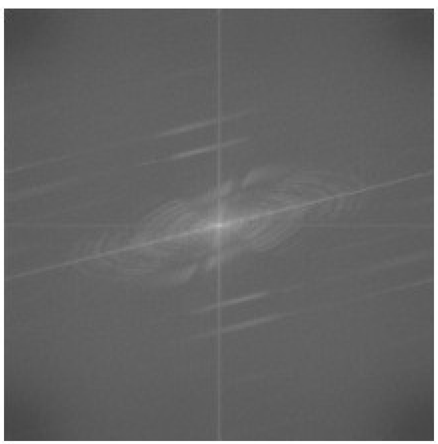
(a) f4.5



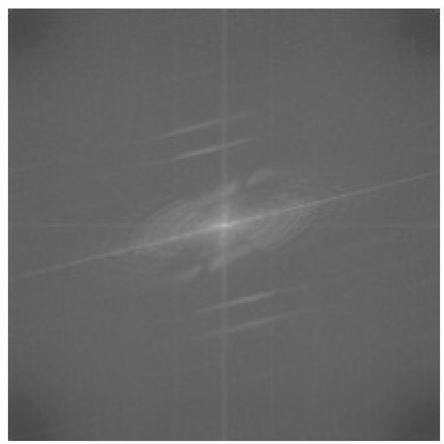
(b) f7.1



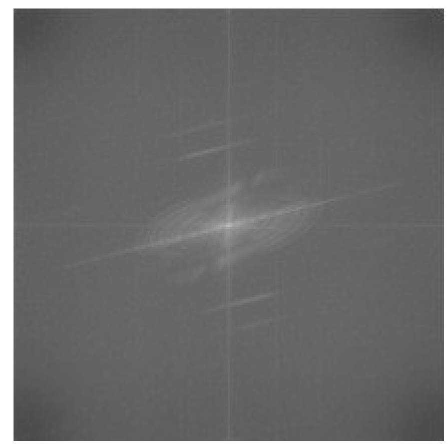
(c) f11



(d) f18



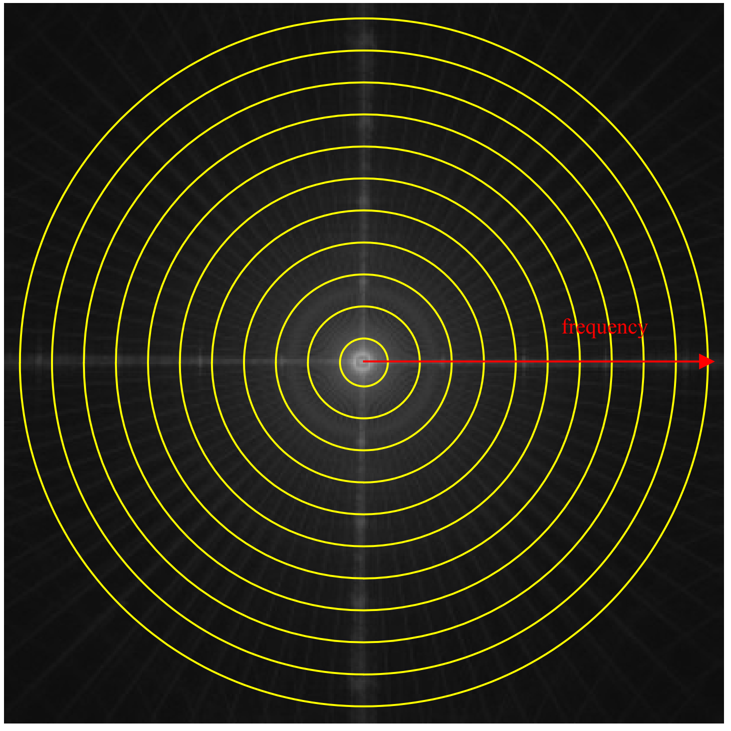
(e) f29

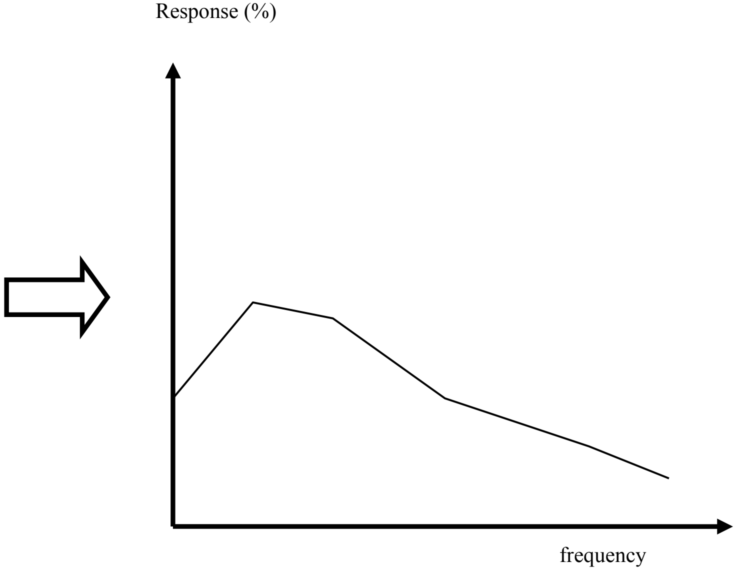


(f) f40

Frequency based analysis

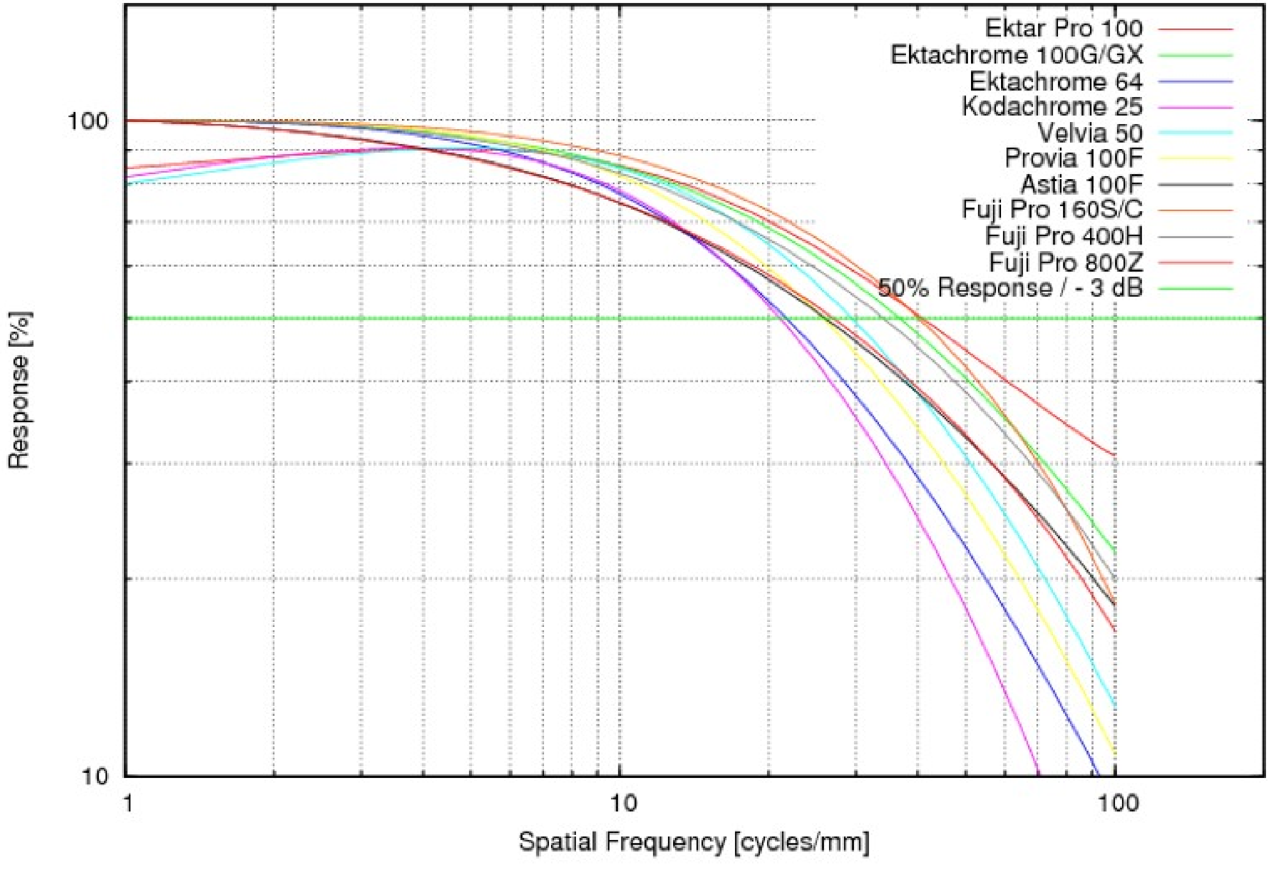
- To see behaviour of the optical system with respect to the different frequency ranges, the FFT values are summed on disks, showing how much different frequency ranges are present in the system





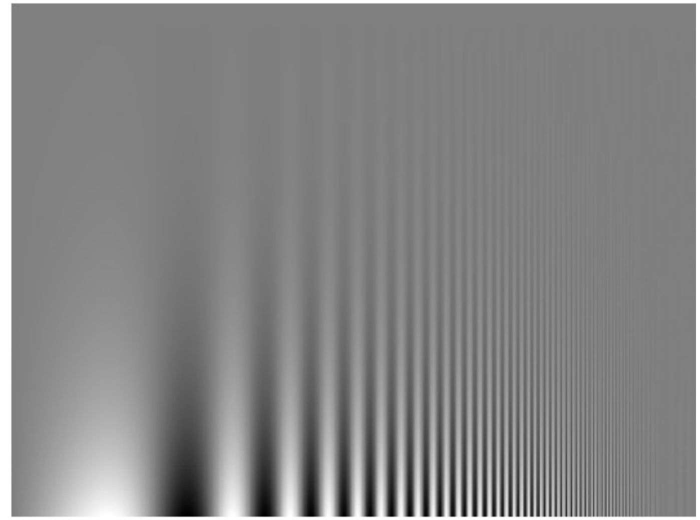
Modulation transfer functions

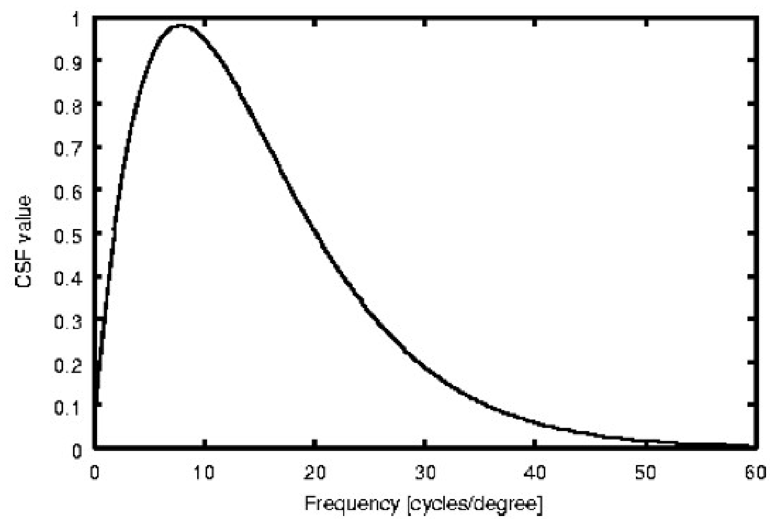
- The response of the optical system is also mediated by the Modulation Transfer Function of the system, which tells how sensitive the optical system is to different frequencies



Contrast Sensitivity Function

- Campbell and Robson did an experiment to determine sensitivity of the human visual system with respect to spacial frequency
- The result is the Constrast sensitivity function pictured below
- It can be used to weigh the analysis in the Frequency domain





The tools: deriving the CWT

- For the Discrete Wavelet Transform, the Fourier basis functions are replaced with wavelets:

$$x(t) = \sum_{n=-\infty}^{\infty} c(n)\phi(t-n) + \sum_{j=\infty}^{\infty}\sum_{n=-\infty}^{\infty} d(j,n)2^{\frac{j}{2}}\psi(2^j t - n)$$

  – stretched and shifted versions of a bandpass function ψ(t)
  – Combined with shifts of a low pass scaling function ϕ(t)
- Coefficients are:
  – Scaling $c(n) = \int_{-\infty}^{\infty} x(t)\phi(t-n)dt$
  – Wavelet $d(j,n) = 2^{j/2}\int_{-\infty}^{\infty} x(t)\psi(2^j t - n)dt$
- DWT:
  – Not shift invariant
  – Does not allow arbitrary orientations

The tools: deriving the CWT

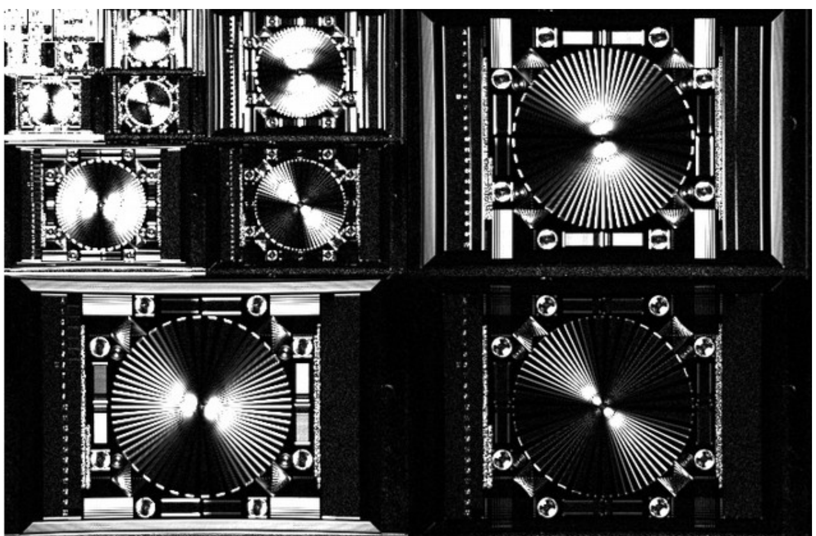
- For the CWT, wavelets in complex form are used:

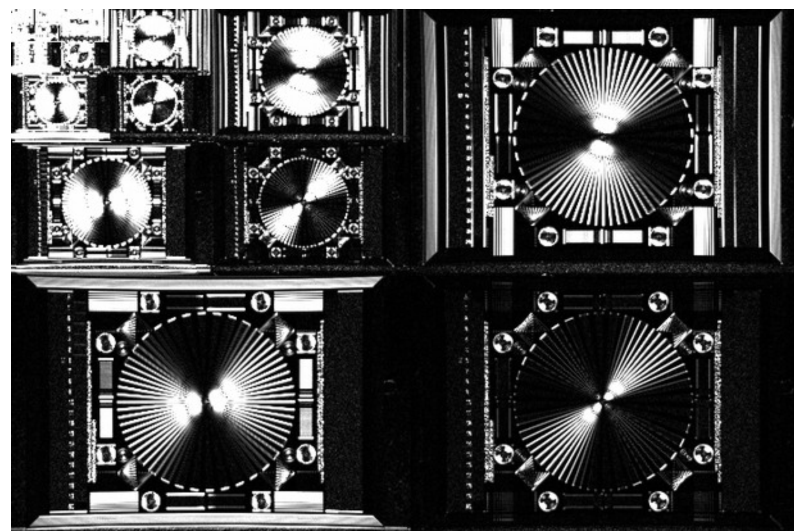
$$\Psi_c(t) = \Psi_r(t) + j\Psi_i(t)$$

  - ψ_r: real part, usually odd
  - ψ_i: imaginary part, usually even
- ψ_r and ψ_i:
  - Can be constructed as a Hilbert pair of *cos(t)* and *sin(t)*.
  - Shifted in phase 90 degrees
  - Computable through Dual Tree Complex Wavelet Transform (*Selesnick et Al.'05*)
  - Coefficients:

$$d_c(j,n) = d_r(j,n) + jd_i(j,n)$$

Subbands:





Use of the CWT

- Like on the FFT, we can observe the distribution of energy with spatial frequency on the CWT.
- CWT allows analyzing sampling in different directions:
  - Computational cost depends on how many directions
- Mean value of the CWT proportional to:
  - Contrast
  - Amount of detail
- It is quite easy to use the CWT for measuring things:
  - Recall: PSNR can be used to measure noise of image

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

  - Take CWT magnitude, replace in PSNR:
    - dynamic range ≈ mean value
    - MSE ≈ standard deviation
  - one obtains Energy Efficiency

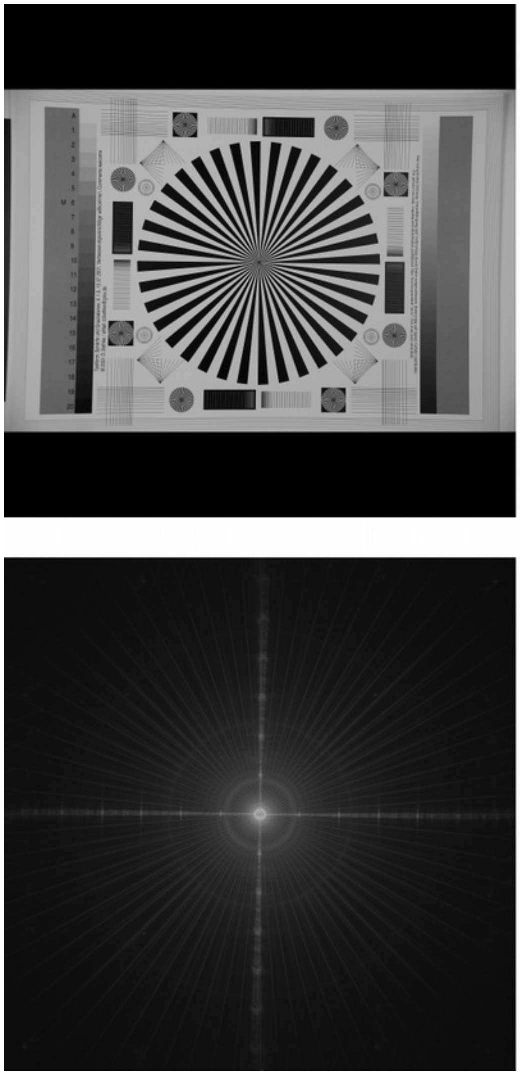
$$EE = \frac{\sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2}}{\bar{x}}$$

- EE measures how much energy of an image is transformed into perceivable detail:
  - Low value ⇒ well distributed detail in the subband (≈ good quality)

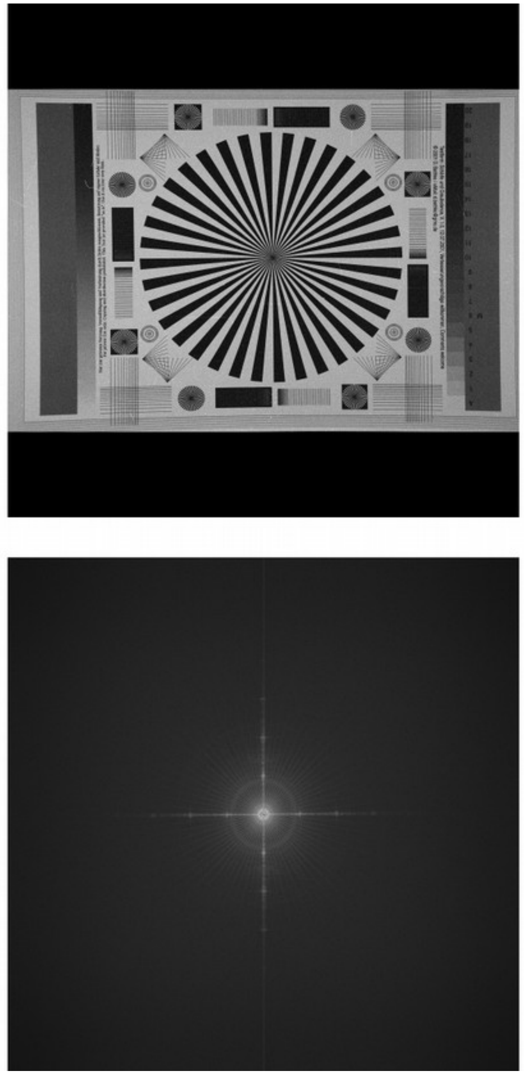
Applications to digital and analogue images

- Let us see if testing the methods on real examples can help in any way:
  - Compare analogue and digital images
  - Pictures taken with same optics: Tamron 28-250mm AF
  - Cameras:
    - Canon 5D Mark II: Sensor size 35mm format (24x36), 16M sensor
    - Canon Analogue Film camera: EOS 650
  - Film used:
    - Kodak TMAX 400
    - Ilford Delta 400
  - Developers used: as specified by manufacturer

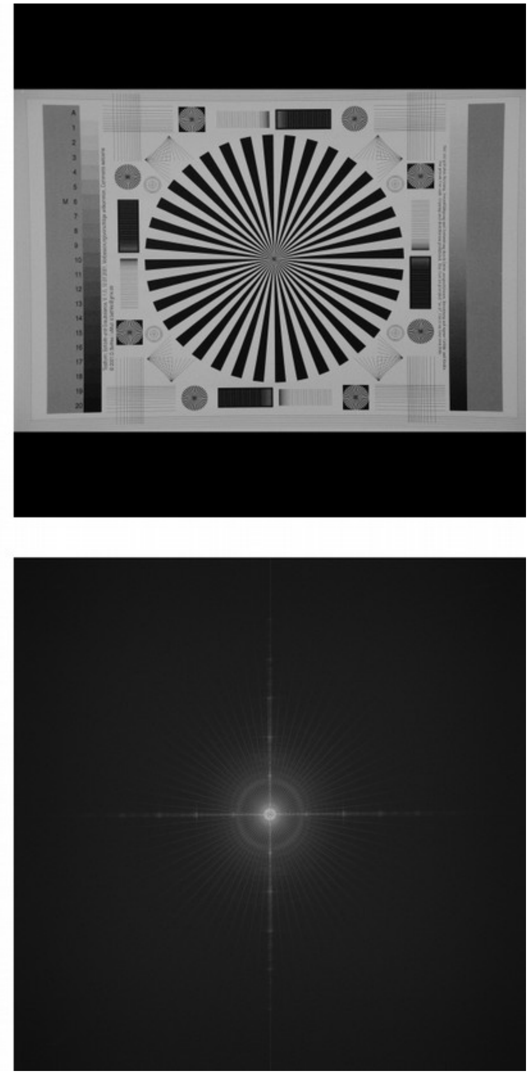
Results: analog vs. digital, FFT



Canon EOS 5D MkII, ISO 400

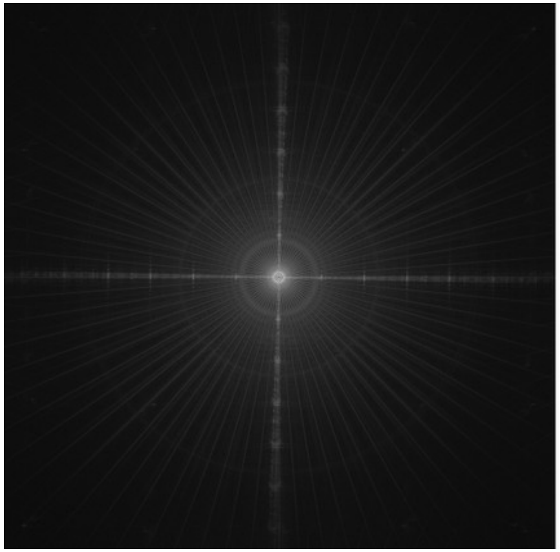


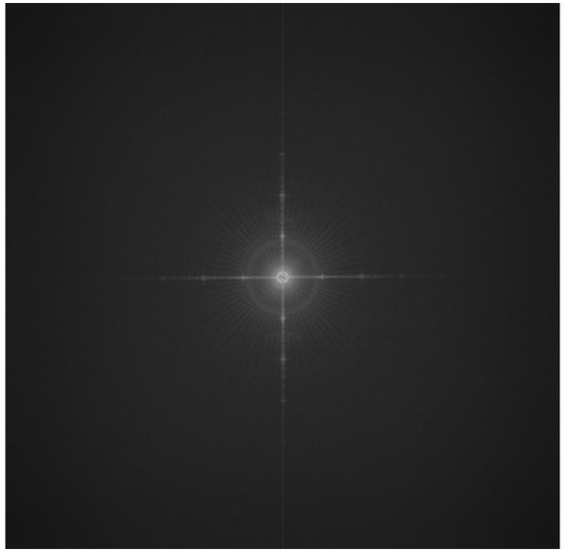
Kodak T-Max ISO 400

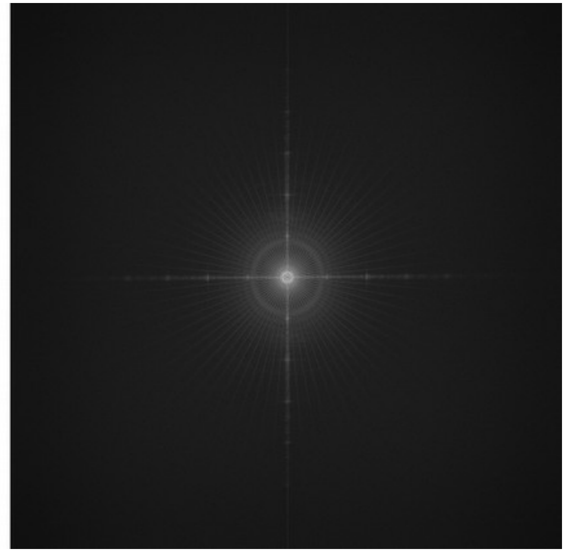


Ilford Delta ISO 400

Results: analog vs. digital, FFT

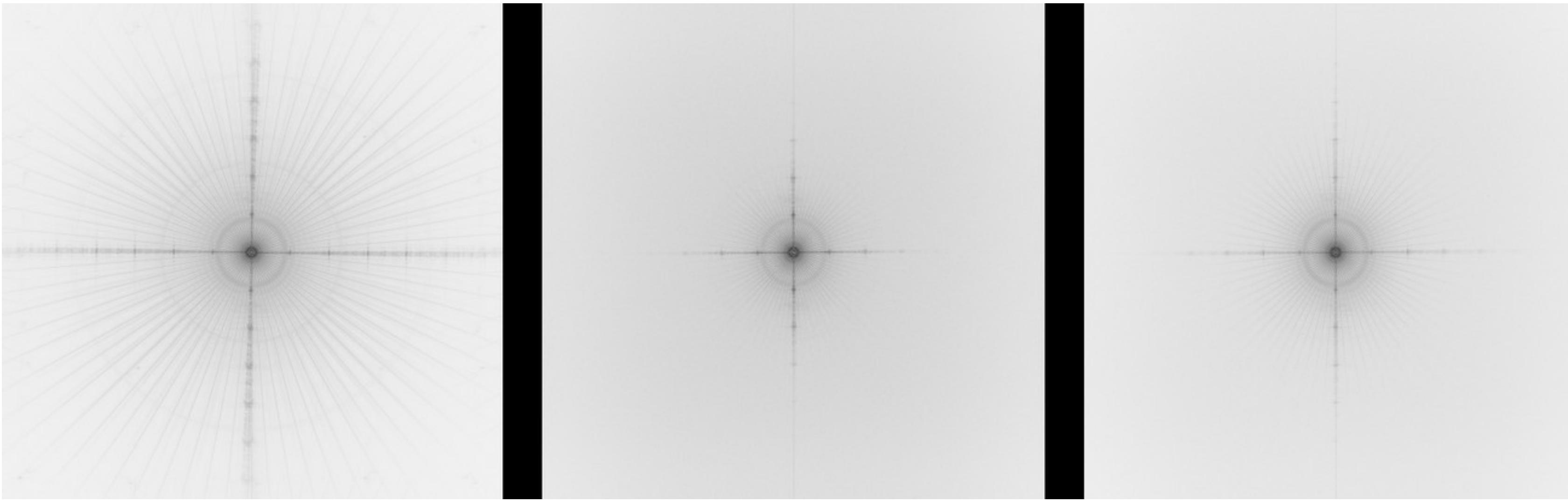






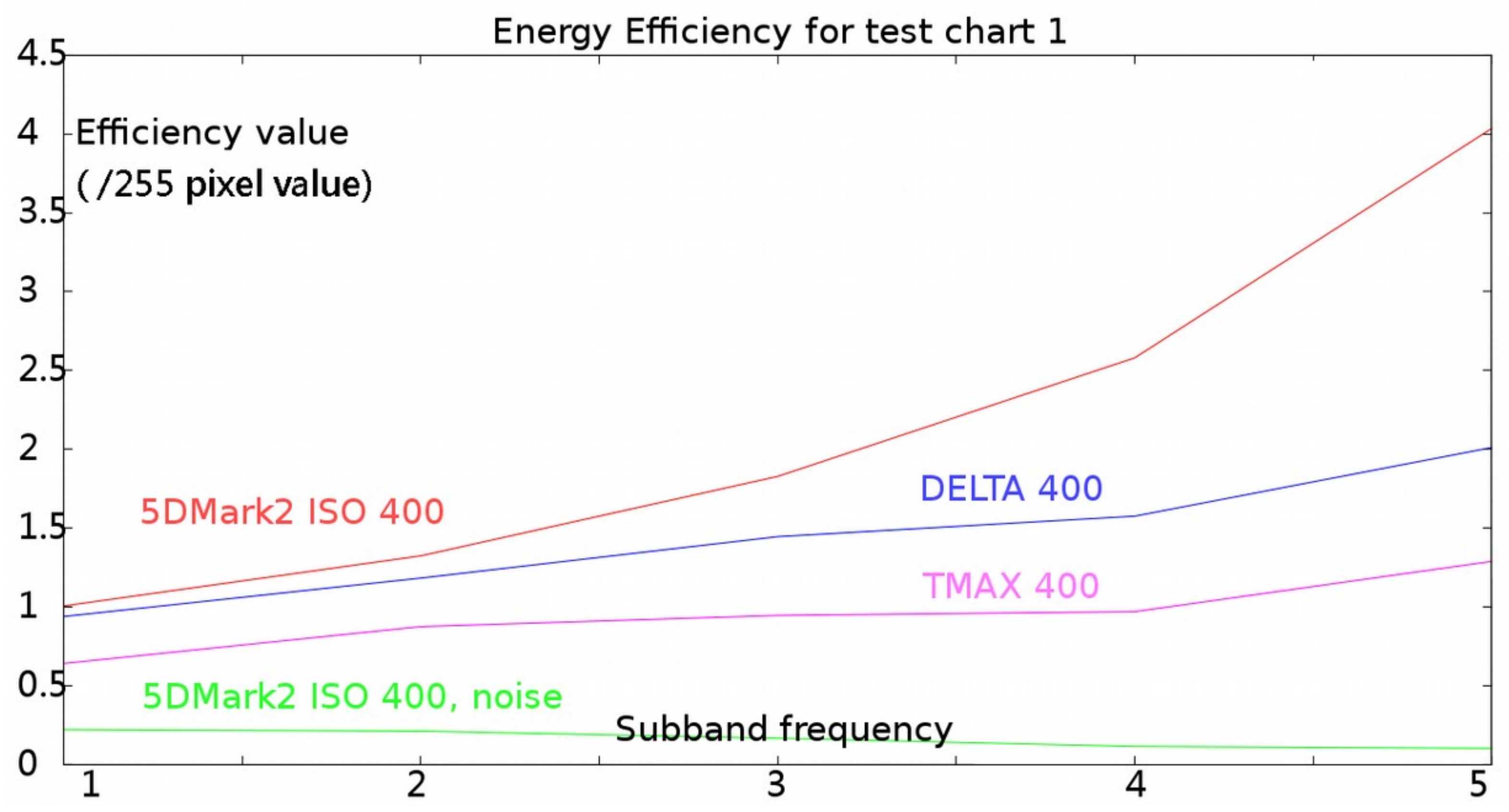
- Canon 5D:
  - Detail over whole spectrum
- T-MAX 400:
  - Cutoff at high frequencies
  - Graininess corrupts detail
  - Perceptual impression: visually not disturbing
- Delta 400:
  - Cutoff at high frequencies
  - Less intense high frequencies

Results: analog vs. digital, FFT



- Canon 5D:
  - Detail over whole spectrum
- T-MAX 400:
  - Cutoff at high frequencies
  - Graininess corrupts detail
  - Perceptual impression: visually not disturbing
- Delta 400:
  - Cutoff at high frequencies
  - Less intense high frequencies

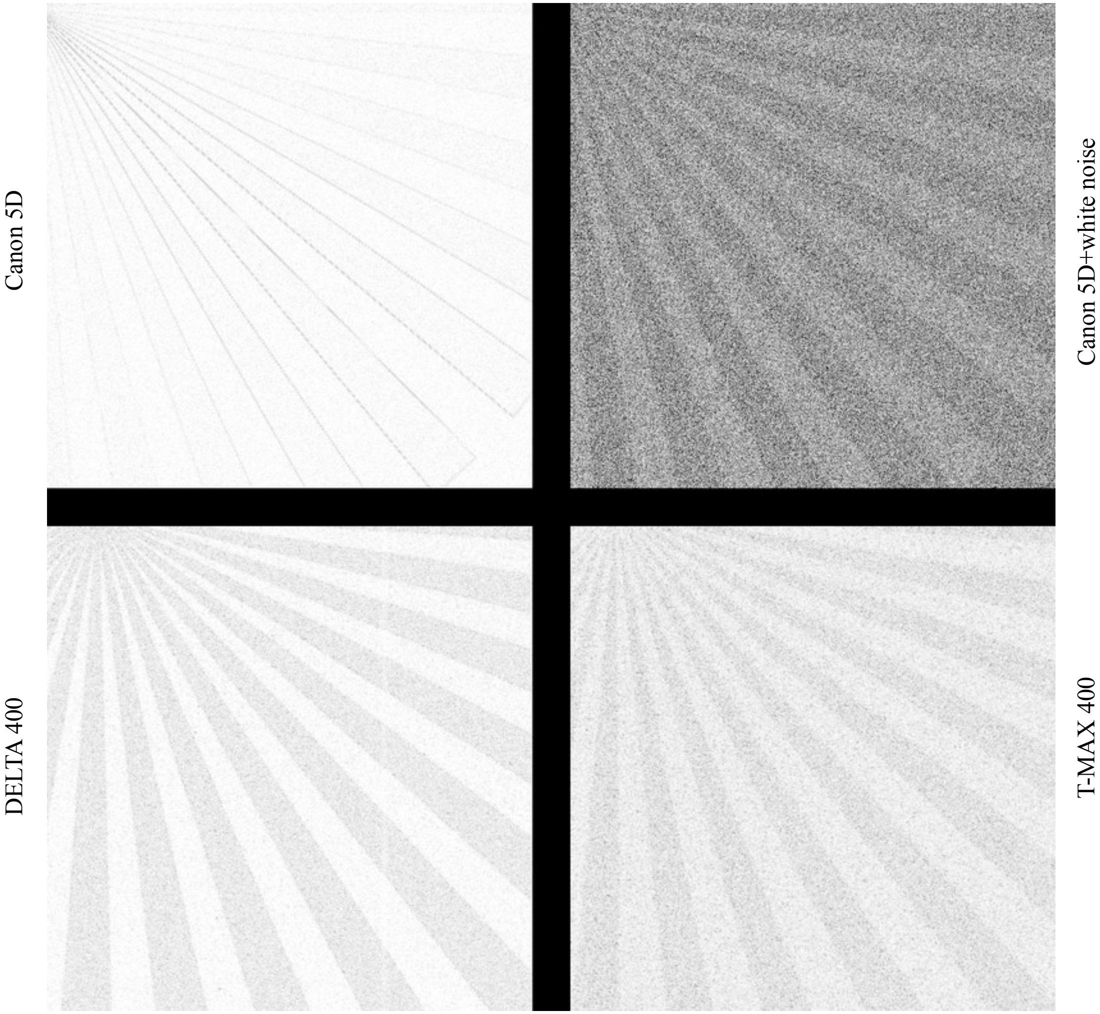
Results: analog vs. digital, CWT-EE



- We additionally added white noise to digital image.
- Reminder: higher=better

- Higher EE of digital camera: this means lower quality.
- But why?

Results: analog vs. digital, CWT-EE

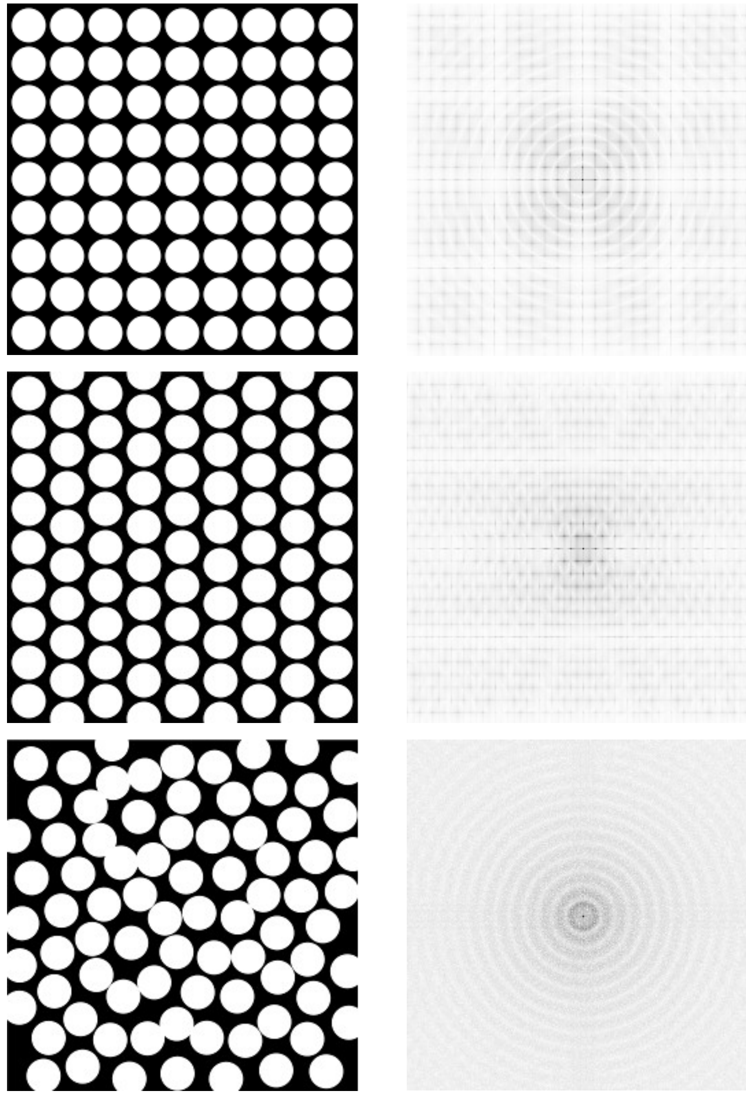


Inverted CWT, highest frequency subband

- Digital camera captures only abrupt changes: edges, contours
  - ⇒ Higher values of EE
- Film captures noisier but smoother changes
  - ⇒ Lower values of the EE
- Notice similar behaviour of films:
  - T-MAX noisier but with better details (upper left corner)

Results: sampling grids and FFT

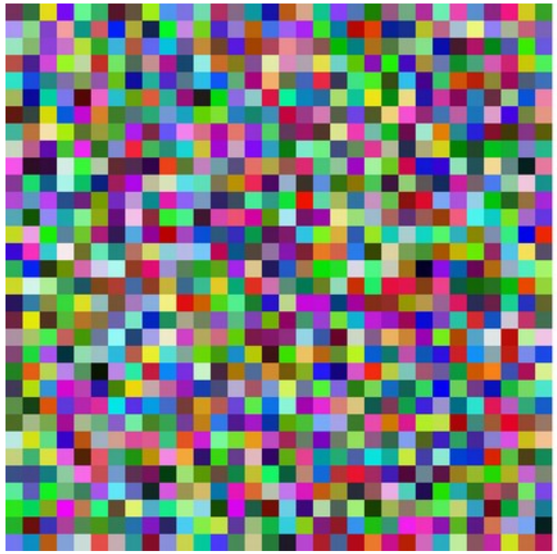
original

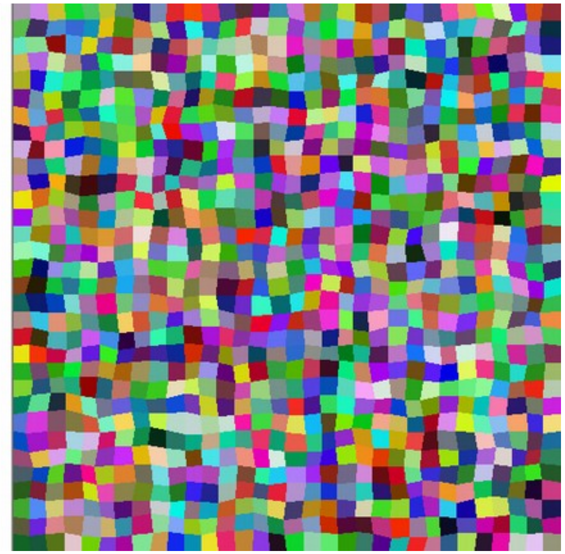


inverted FFT

Resampling pictures at low resolution

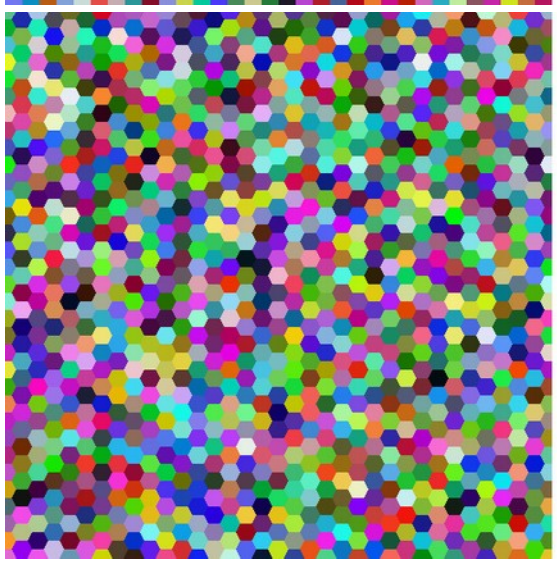
Square

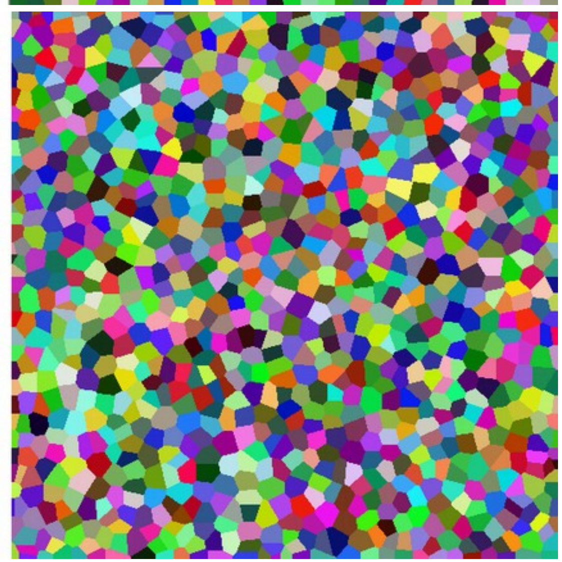




Jittered

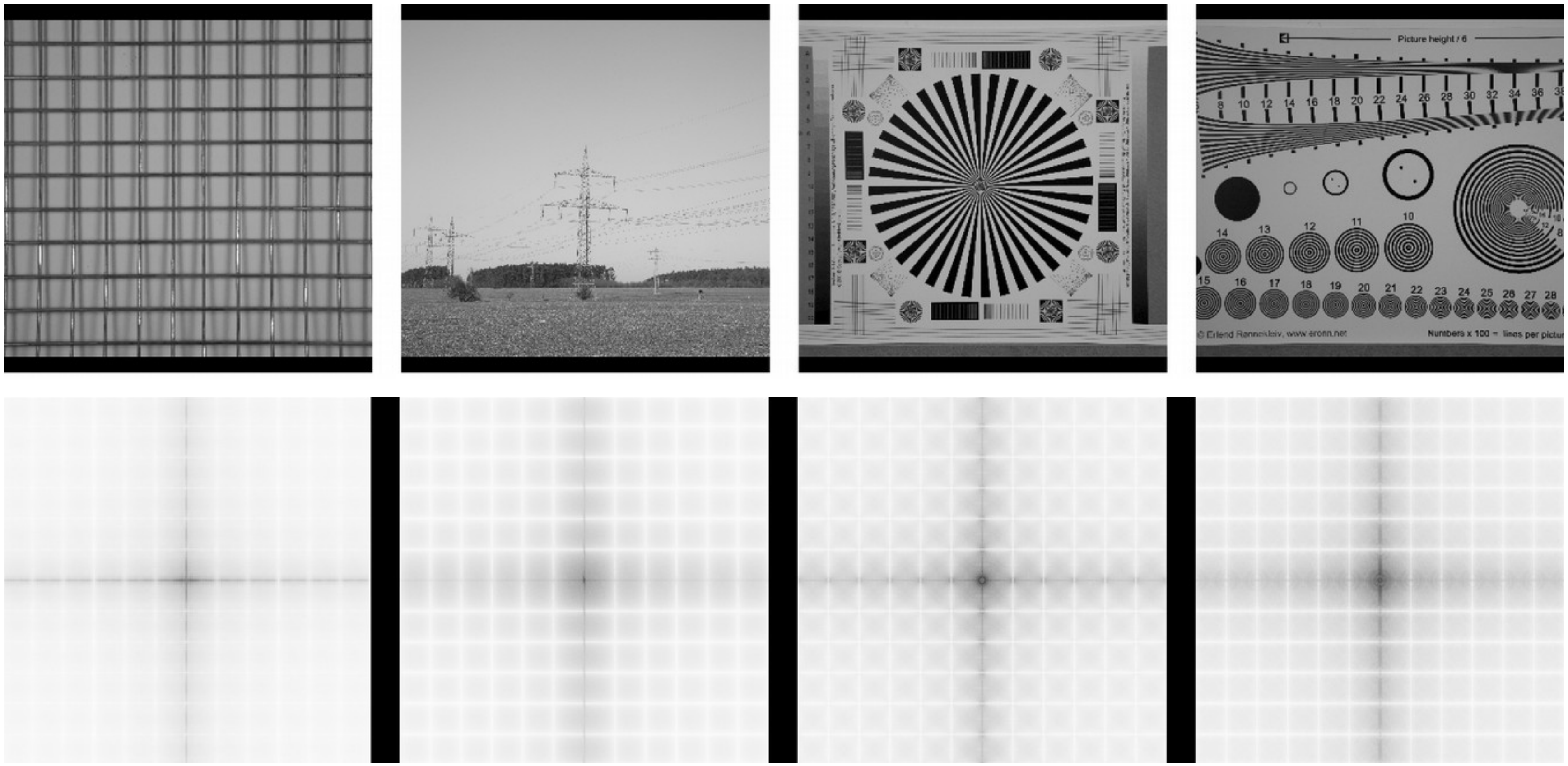
Hexagonal





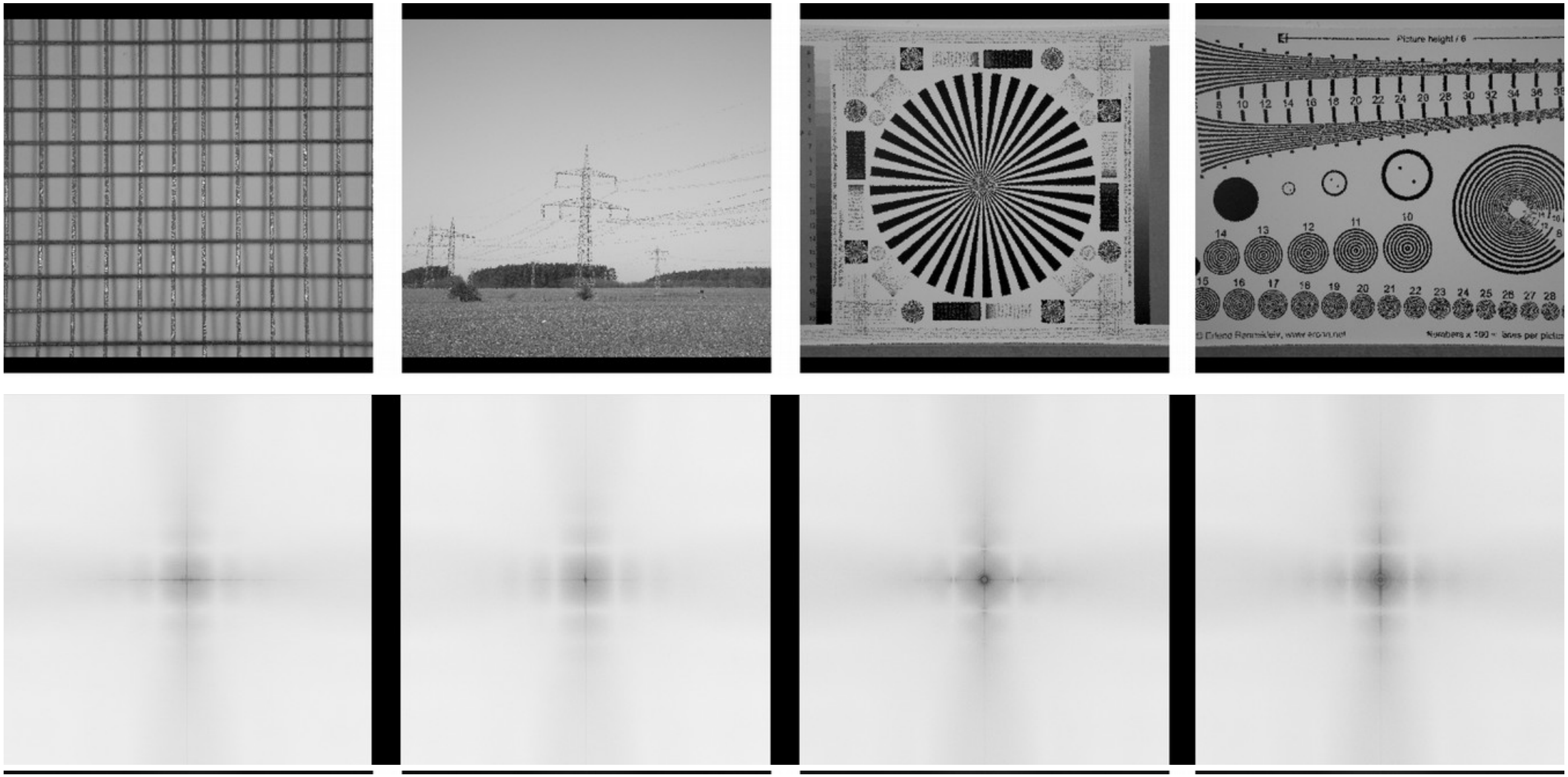
Random

Resampled images: FFT, square



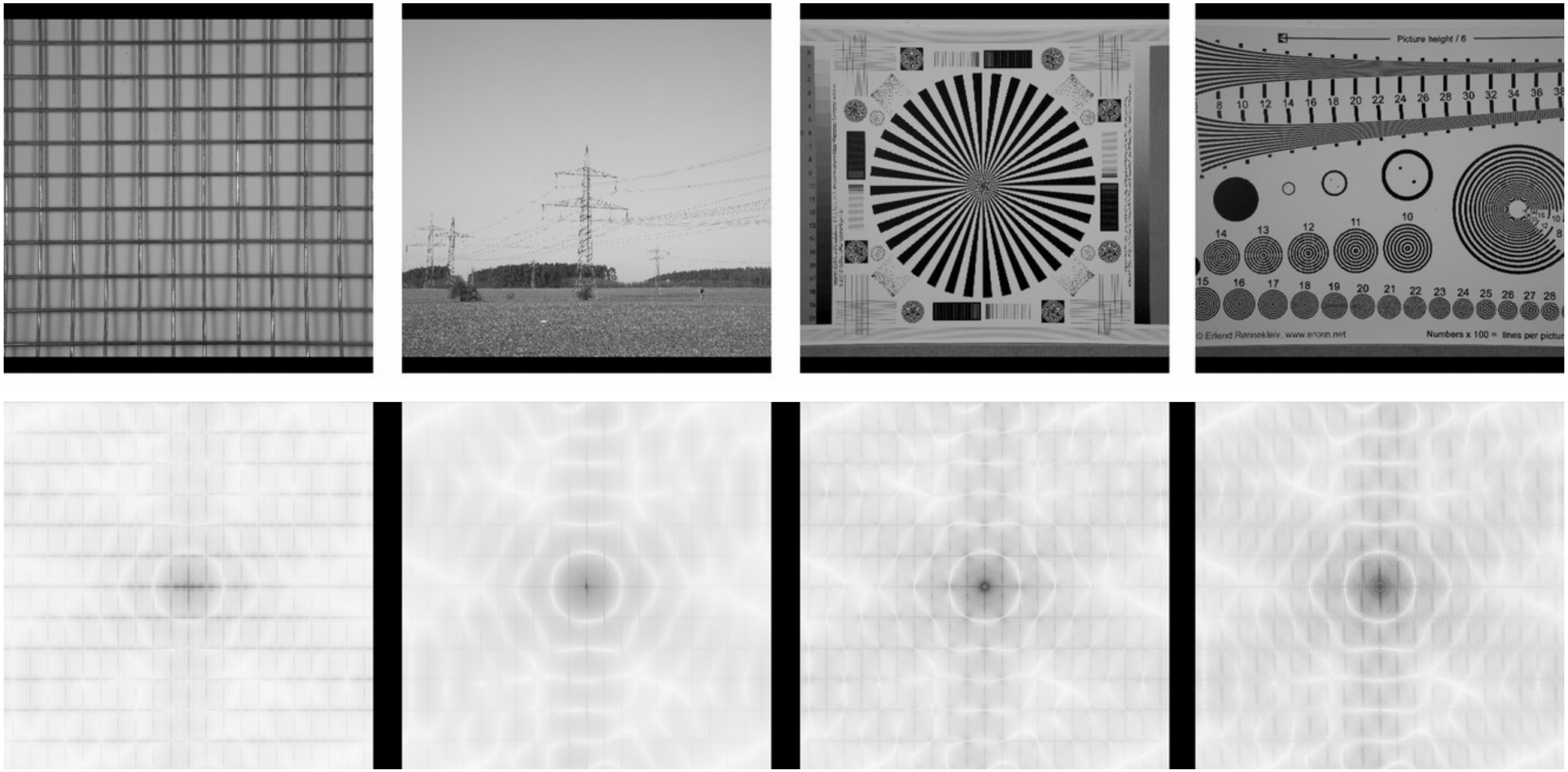
- Resampled: pixel=16x16 pixels
  ca. 100K squares
- Visible base spectrum+replicas
- Aliasing=replica overlap
- Most energy on x and y axis
- Powerline strongly aliased

Resampled images: FFT, jittered



- Same size as square.
- Energy still distributed along axes
- Grid fades at high frequencies
- Much less overlaps:
  less aliasing!

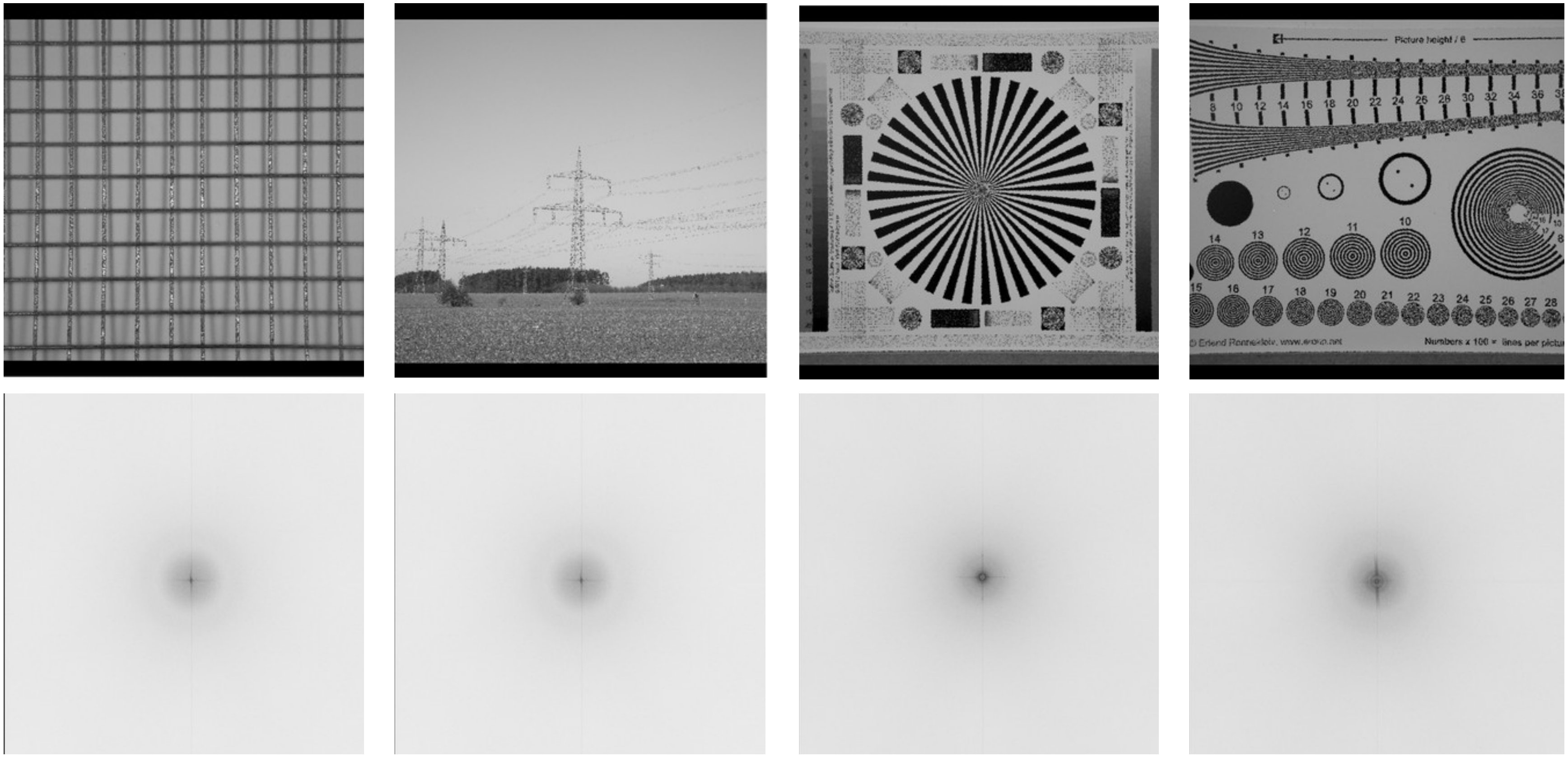
Resampled images: FFT, hexagonal



- Side size: 7 pixels
  120K pixel per picture
- 3 axes: x axis, 30^o and -30^o

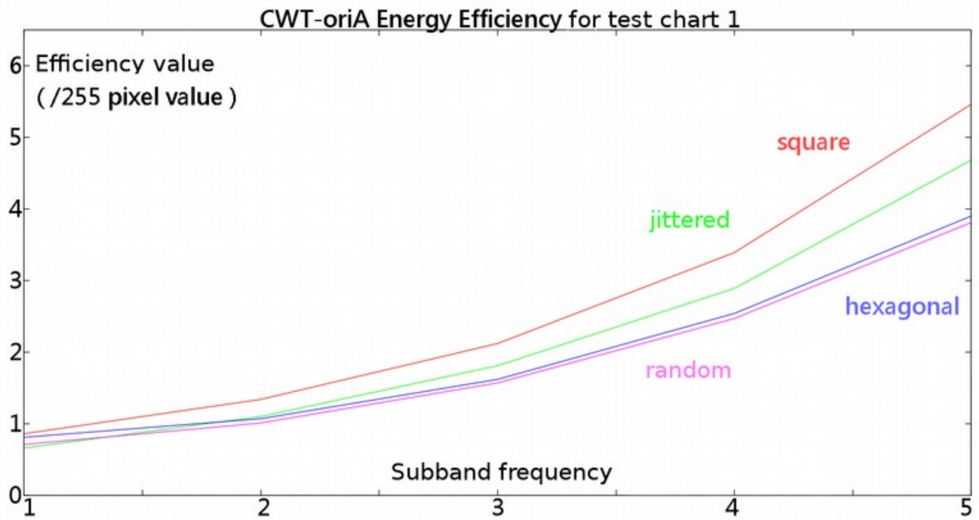
- Repeating hexagons
- More aliasing than jittering
- Decent compromise for hardware

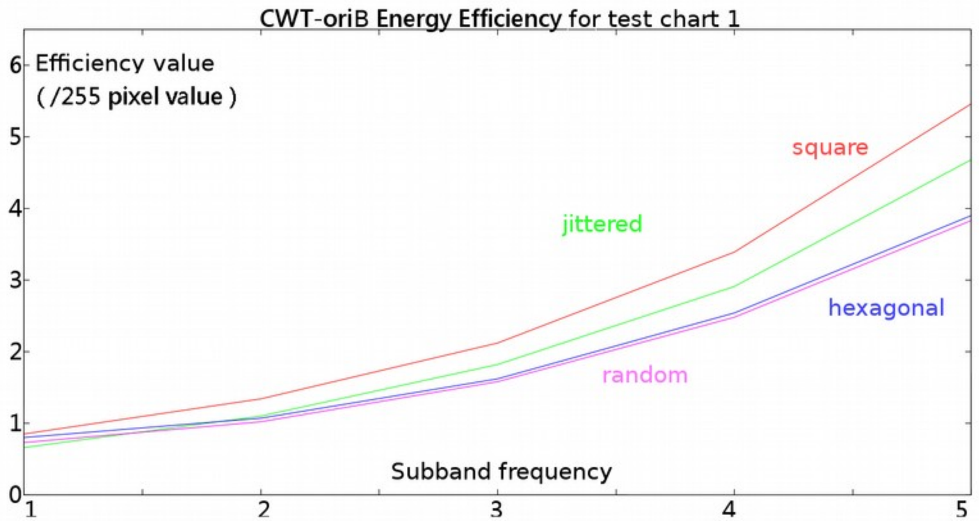
Resampled images: FFT, random Voronoi



- Voronoi sets from random numbers: 160K points
- No energy concentration on any axis.
- No replicas either: just „base“ spectrum
- No aliasing: optimal sampling!

Resampled images: CWT-EE





- EE: similar characteristics for the six directions of the CWT
  - Square: higher EE
  - Random: lowest
  - Jittered: inbetween
  - Hexagonal: almost like random
- Little difference at small frequency
- Interesting: similar to what we found out for digital sensors

Conclusions on FFT and CWT

- Using FFT and EE of the CWT for image analysis can provide useful information on image quality
  - FFT: directions of sampling, antialiasing, presence of detail.
  - CWT: allows inspection of frequency subbands along different directions
- We can conclude:
  - Randomized sampling behaves better since it samples all directions equally
  - Disappearing details in digital pictures are due to square grid sampling, which favours certain directions
  - Analog film does not have the same sharpness in detail, but
    - It samples better
    - It is more pleasant to the eye

Quality: the higher level

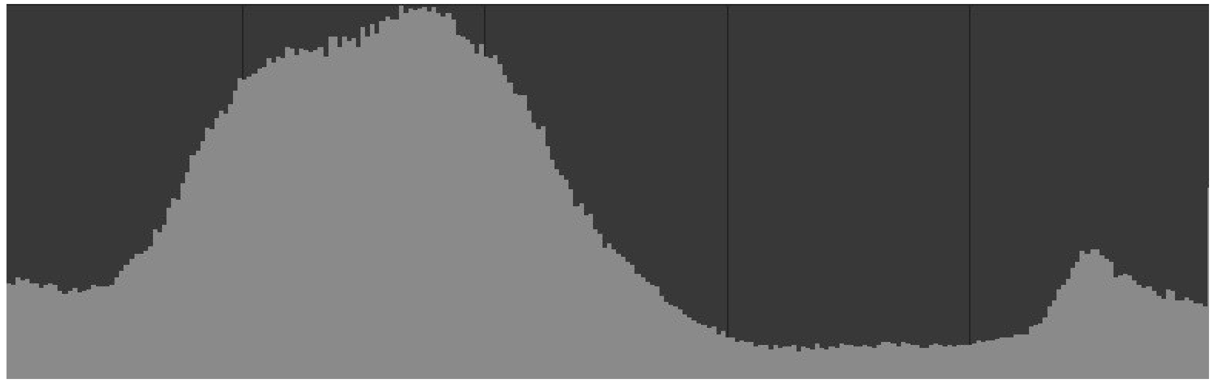
- We move now from micro to the macro level
- This is work in progress, and therefore just a couple of tentative research directions
- The goal: identify good characteristics of images

- Which rules do photographers use?

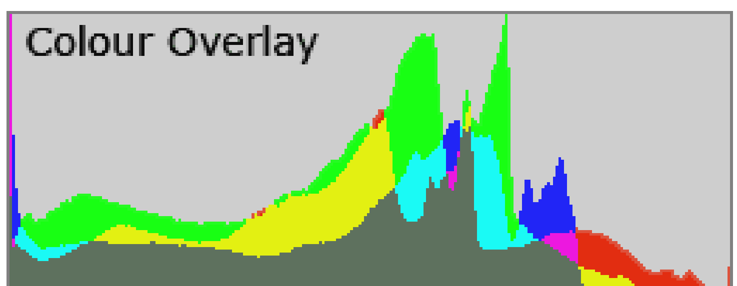
- Photography is not as simple as it looks like
  - Plenty of rules
  - Lots of guidelines
  - Fine craftmanship which need:
    - Eye for subject
    - Knowledge of rules
    - Knowledge of light
    - Knowledge of colour
    - Knowledge of composition
    - Knowledge of camera, shutter speed, aperture valeus

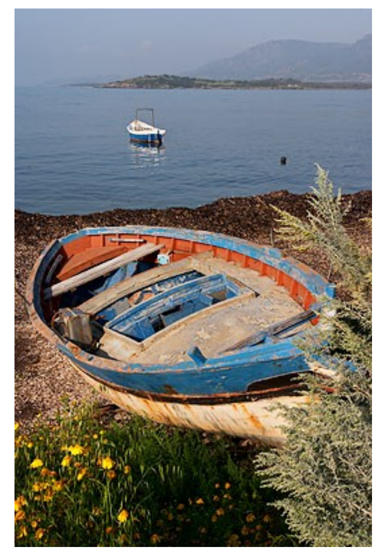
Making computable sense of rules: light - colour

- Good lighting can be seen on the greyscale histogram



- Color rendition on simple RGB diagrams can be seen similarly…
- Although in color Hues should be more important



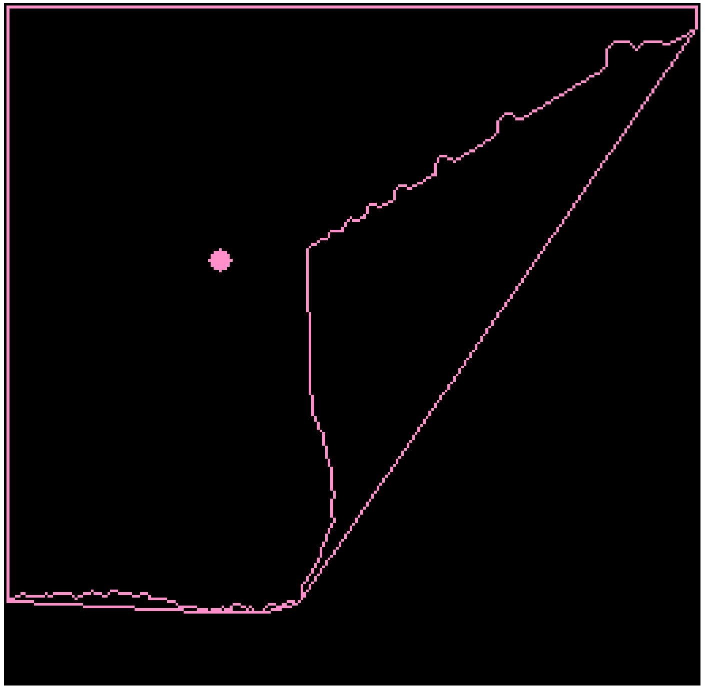


Composition: automatic rule of thirds

- We started from a "simple" rule: the rule of thirds
- Run contour detection on a simple image and detect connected components
- Compute convex hull of contours
- Compute centers of convex hulls: useful for rule of thirds (marked points).

Composition: automatic rule of thirds

- However, the convex hull is tricky:
  is the subject light or dark?

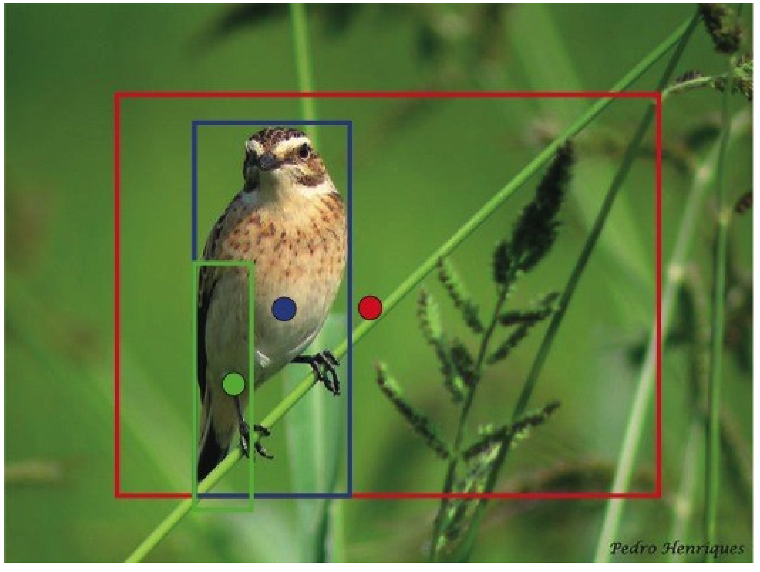


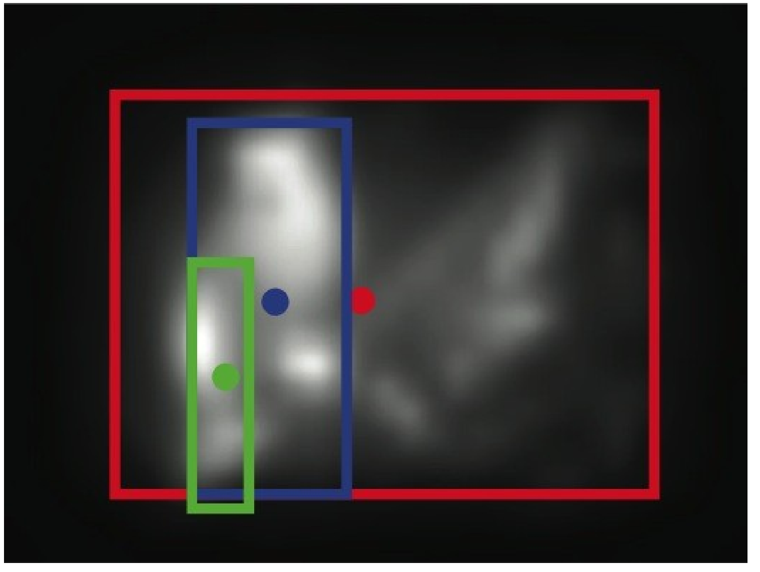
Inverted background/foreground:

incorrect convex hull and center!

Composition: automatic rule of thirds

- Recently, machine learning techniques have been used for the detection of the rule of thirds
- In the literature, some authors use *saliency maps* for the same purpose (Mai, Le, Niu and Liu 2011)

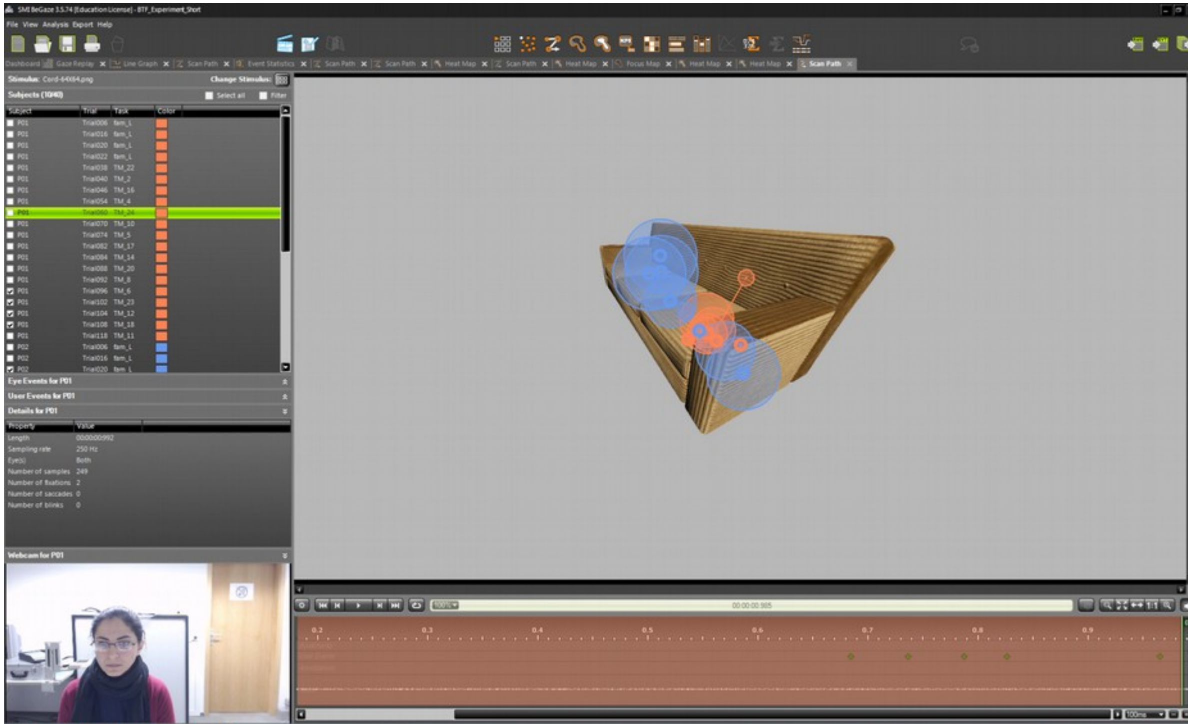


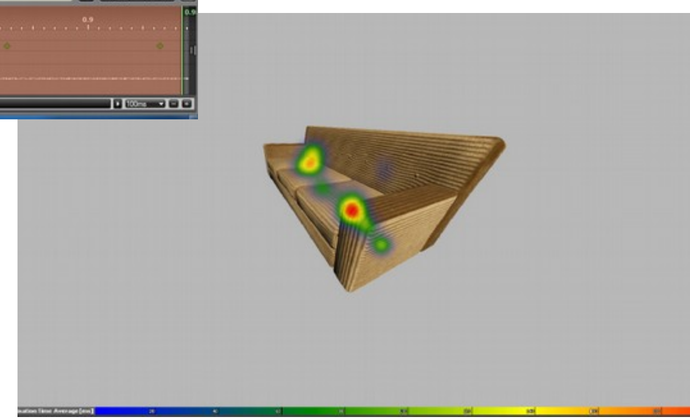


Perception and Image Quality

- Measuring gaze in humans when perceiving high quality renderings (BTF)
- Psychophysical Experiments using eye trackers
- Perceptual Evaluation of rendering quality

Perception and Image Quality





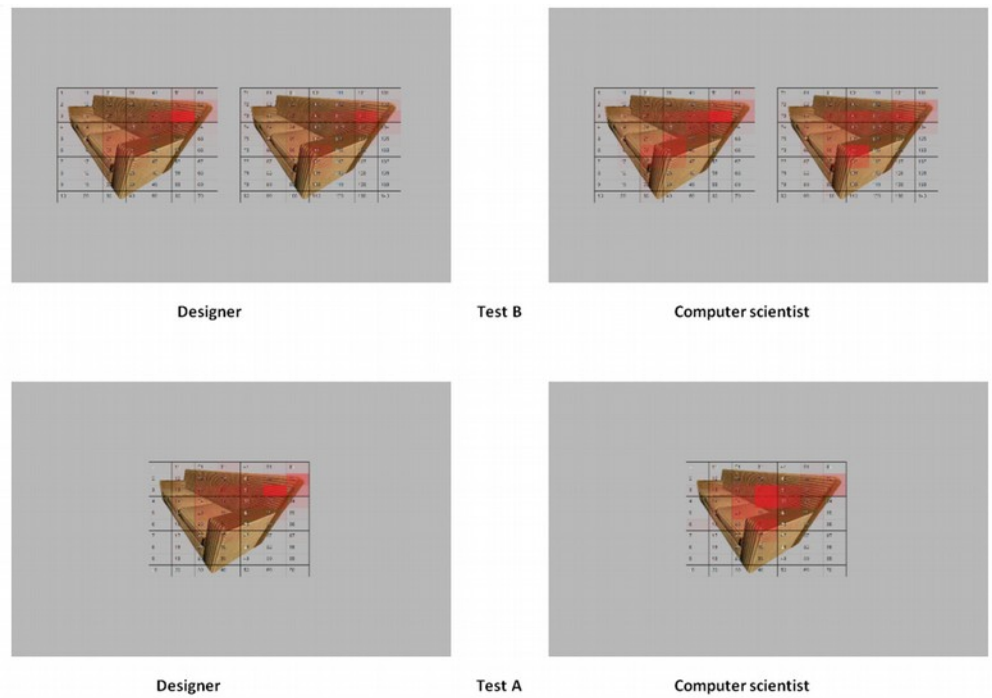
Different texture resolutions:

does the user notice?

Perception and Image Quality

- First result: the performance of artists in perceiving quality is significantly better than the performance of computer scientists

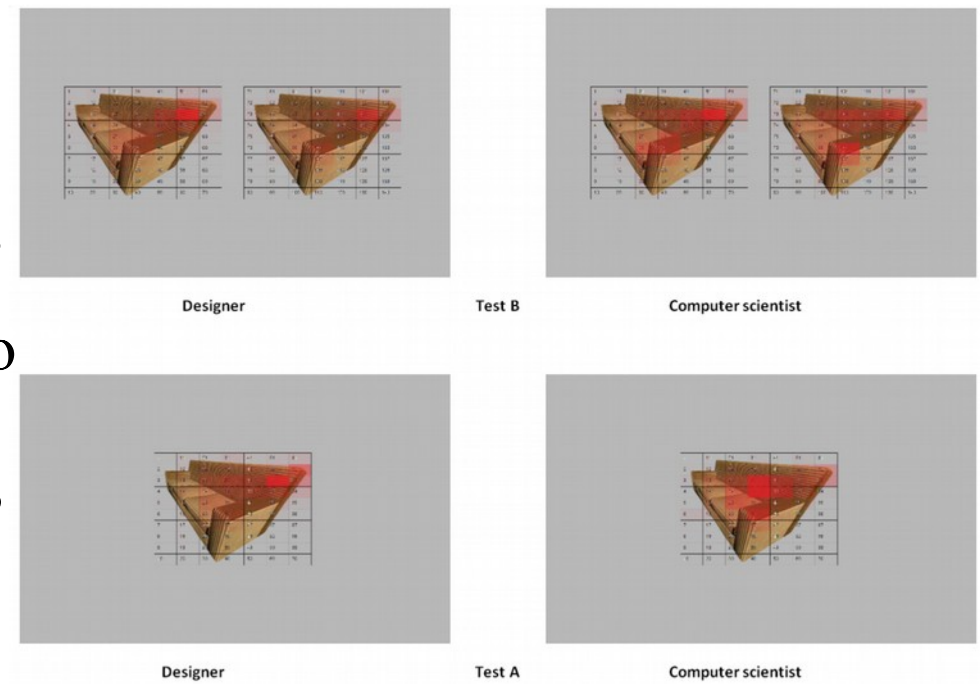
	Test B		Test A	
	D	CS	D	CS
Mean	24.43	20.19	36.10	32.19
Std.	4.422	6.071	12.739	14.784
Median	25.00	21.00	36.00	32.00
Normal distribution	yes		yes	
Significance	t(40)=2.586, p<.05		t(40)=.917, p>.05	
Accepted	yes		No	



Perception and Image Quality

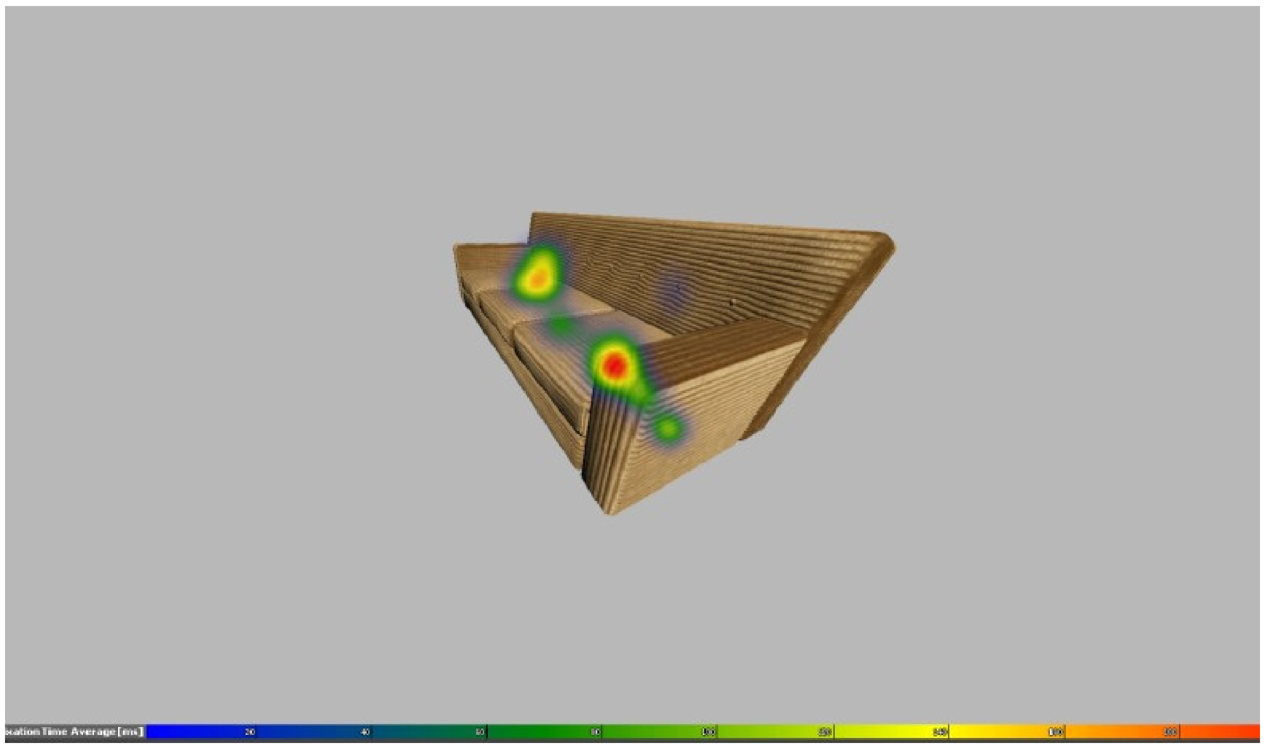
- Relevant questions: how long does it take to detect differences?
- Do different skills matter?
- What is the time it takes to detect differences?
- Are there gazing patterns?

- Ultimately: can I cheat on rendering quality without users noticing?



Perception and Image Quality

- Second result:
  Users do not notice differences if the texture has a resolution bigger than 128x128

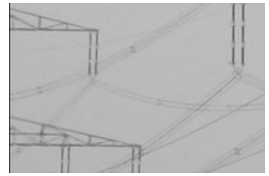
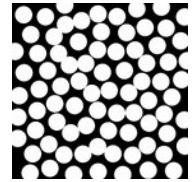


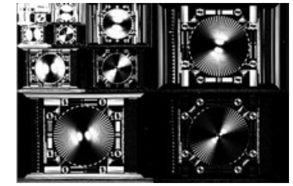
Future work

- We just took a bite at a HUGE cheese
- Open issues:
  - Conceive a “standard” sample set which makes sense and is shot under controlled conditions.
  - Explore possible measures outlining FFT characteristics: one number for directionality.
  - Include knowledge on human visual system.
  - What happens if I rotate CWT directions?
    - Varying EE? Is there a linking function?
  - Open consequences for sensors and displays.
  - Continue experimenting with perception and quality

Acknowledgements and coordinates





- Credits:
  - Photo samples author: Igor Ziesche
  - Part of the illustrations and picture material can be found in Chen Jiayin's Masters Thesis: "Image Quality Evalutation using Structural Similarity Index Maps", Faculty of Media, Bauhaus-University Weimar, 2011
  - Study on human perception of images in Ph.D. thesis of Dr. Banafsheh Azari: „Bidirectional Texture Functions: Acquisition, Rendering and Quality Evaluation", Fakultät Medien, Bauhaus-Universität Weimar, Mai 2018.

- Hope you liked the course:
  thank you for your patience …err attention!

- More info:
  http://www.uni-weimar.de/medien/cg