Autofocusing with the help of the **empirical Haar transform**

Przemysław Śliwiński and Krzysztof Berezowski

Institute of Computer Engineering, Control and Robotics
Wrocław University of Technology, POLAND

WASC 2012, Clermont-Ferrand, April 5-6th, **2012**

Introduction

Presentation schedule

- Motivations and inspirations

Introduction

Presentation schedule

- Motivations and inspirations
- Model and formal assumptions

Introduction

Presentation schedule

- Motivations and inspirations
- Model and formal assumptions
- Generic algorithm and its properties

Introduction

Presentation schedule

- Motivations and inspirations
- Model and formal assumptions
- Generic algorithm and its properties
- AF criteria

Introduction

Presentation schedule

- Motivations and inspirations
- Model and formal assumptions
- Generic algorithm and its properties
- AF criteria
- Unbalanced Haar Transform and Single-Photon AF

Introduction

Presentation schedule

- Motivations and inspirations
- Model and formal assumptions
- Generic algorithm and its properties
- AF criteria
- Unbalanced Haar Transform and Single-Photon AF
- Experimental results and conclusions

Introduction

Motivations and inspirations

Problem

A proper and reliable focusing algorithm is a conditio sine qua non of a 'good image'. Not only from an aesthetic vantage point, but also in automated applications.

- We exploit a *plethora* of the 'off-the-shelf' theoretical results developed in various disciplines:

Introduction

Motivations and inspirations

Problem

A proper and reliable focusing algorithm is a conditio sine qua non of a 'good image'. Not only from an aesthetic vantage point, but also in automated applications.

- We exploit a *plethora* of the 'off-the-shelf' theoretical results developed in various disciplines:
  - *signal and image processing*, *image analysis*, *harmonic analysis*, *control theory*, or

Introduction

Motivations and inspirations

Problem

A proper and reliable focusing algorithm is a conditio sine qua non of a 'good image'. Not only from an aesthetic vantage point, but also in automated applications.

- We exploit a *plethora* of the 'off-the-shelf' theoretical results developed in various disciplines:
  - *signal and image processing*, *image analysis*, *harmonic analysis*, *control theory*, or
  - *information theory*, *probability theory* and *mathematical statistics*, as well.

Introduction
Alternatives

- **Stereo-vision**

Solution

Our algorithm works with standard matrix sensors & standard optics, and employs standard transforms and routines...

Introduction
Alternatives

- **Stereo-vision**
  - two sensors

Solution

Our algorithm works with standard matrix sensors & standard optics, and employs standard transforms and routines. . .

Introduction

Alternatives

- **Stereo-vision**
  - two sensors
  - two lenses, etc.

Solution

Our algorithm works with standard matrix sensors & standard optics, and employs standard transforms and routines...

Introduction

Alternatives

- **Stereo-vision**
  - two sensors
  - two lenses, etc.
- **Light-field cameras**

Solution

Our algorithm works with standard matrix sensors & standard optics, and employs standard transforms and routines. . .

Introduction

Alternatives

- **Stereo-vision**
  - two sensors
  - two lenses, etc.
- **Light-field cameras**
  - lack resolution/dynamic range

Solution

Our algorithm works with standard matrix sensors & standard optics, and employs standard transforms and routines...

Introduction

Alternatives

- **Stereo-vision**
  - two sensors
  - two lenses, etc.
- **Light-field cameras**
  - lack resolution/dynamic range
  - *computational photography* devices

Solution

Our algorithm works with standard matrix sensors & standard optics, and employs standard transforms and routines...

Introduction

Alternatives

- **Stereo-vision**
  - two sensors
  - two lenses, etc.
- **Light-field cameras**
  - lack resolution/dynamic range
  - *computational photography* devices
- **Femtosecond lasers**

Solution

Our algorithm works with standard matrix sensors & standard optics, and employs standard transforms and routines...

Introduction
Alternatives

- **Stereo-vision**
  - two sensors
  - two lenses, etc.
- **Light-field cameras**
  - lack resolution/dynamic range
  - *computational photography* devices
- **Femtosecond lasers**
  - comparatively slow (like line scanners)

Solution

Our algorithm works with standard matrix sensors & standard optics, and employs standard transforms and routines. . .

Introduction

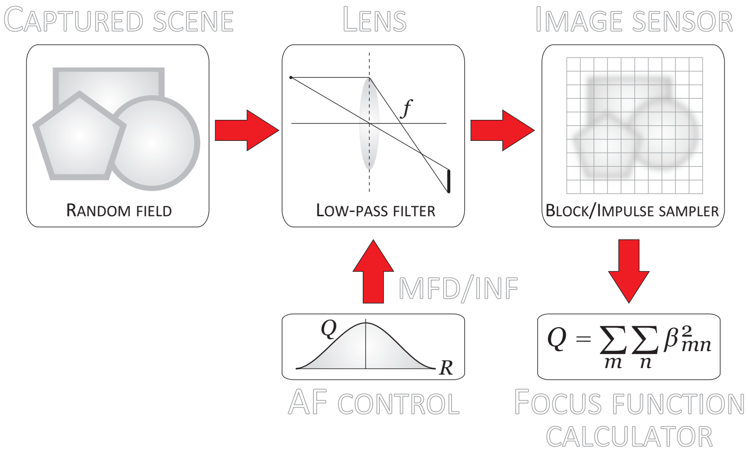
Alternatives

- **Stereo-vision**
  - two sensors
  - two lenses, etc.
- **Light-field cameras**
  - lack resolution/dynamic range
  - *computational photography* devices
- **Femtosecond lasers**
  - comparatively slow (like line scanners)
  - *computational photography* devices

Solution

Our algorithm works with standard matrix sensors & standard optics, and employs standard transforms and routines. . .

Problem statement
AF model



Generic AF algorithm steps

1. Compute the focus function (with optional:

Generic AF algorithm steps

1. Compute the focus function (with optional:
    1. denoising and

Generic AF algorithm steps

1. Compute the focus function (with optional:
    1. denoising and
    2. sensor output linearization).

Generic AF algorithm steps

1. Compute the focus function (with optional:
    1. denoising and
    2. sensor output linearization).
2. Shift the lens accordingly:

Generic AF algorithm steps

1. Compute the focus function (with optional:
   1. denoising and
   2. sensor output linearization).
2. Shift the lens accordingly:
   1. determine the direction

Generic AF algorithm steps

1. Compute the focus function (with optional:
   1. denoising and
   2. sensor output linearization).
2. Shift the lens accordingly:
   1. determine the direction
   2. set the step-size

Generic AF algorithm steps

1. Compute the focus function (with optional:
    1. denoising and
    2. sensor output linearization).
2. Shift the lens accordingly:
    1. determine the direction
    2. set the step-size
3. Make it **reliable** in *noisy environments*!

Problem statement

Assumptions

1. The **scene** is a *2D homogenous second-order stationary process* (thus an *ergodic (in the wide sense) random field*) with unknown distribution and unknown correlation function.

Problem statement

Assumptions

1. The **scene** is a *2D homogenous second-order stationary process* (thus an *ergodic (in the wide sense) random field*) with unknown distribution and unknown correlation function.
2. The **lens** system is modeled with the help of the *first-order optics laws*, that is, the lens is merely a *simple centered moving average* filter with an order proportional to the distance of the sensor from the image plane and to the size of the lens aperture.

Problem statement

Assumptions

1. The **scene** is a *2D homogenous second-order stationary process* (thus an *ergodic (in the wide sense) random field*) with unknown distribution and unknown correlation function.
2. The **lens** system is modeled with the help of the *first-order optics laws*, that is, the lens is merely a *simple centered moving average* filter with an order proportional to the distance of the sensor from the image plane and to the size of the lens aperture.
3. The **image sensor** acts as a *block sampler*, that the lens-produced image is *orthogonally projected* onto the space of piecewise constant functions.

AF algorithm foundations

- The AF algorithm is based on the following lemmas:

Lemma

Under assumptions **1-3**, *the variance of the captured image is a* **unimodal** *function w.r.t. the order of the lens filter and attains its* **maximum** *value for the in-focus image.*

Lemma

The variance estimate is **tantamount** *to the orthogonal expansion of the image acquired by the sensor.*

AF algorithm routine

- The algorithm seeks the **maximum** of the (noised) **image variance**.

AF algorithm routine

- The algorithm seeks the **maximum** of the (noised) **image variance**.
- The focus functions is **global!**

AF algorithm routine

- The algorithm seeks the **maximum** of the (noised) **image variance**.
- The focus functions is **global!**
- Various (bi-)**orthogonal expansions** can be used to estimate the variance:

AF algorithm routine

- The algorithm seeks the **maximum** of the (noised) **image variance**.
- The focus functions is **global!**
- Various (bi-)**orthogonal expansions** can be used to estimate the variance:
  - trigonometric (DCT, *Fourier*, *Hartley*),

AF algorithm routine

- The algorithm seeks the **maximum** of the (noised) **image variance**.
- The focus functions is **global!**
- Various (bi-)**orthogonal expansions** can be used to estimate the variance:
  - trigonometric (DCT, *Fourier*, *Hartley*),
  - *Walsh-Hadamard* (additions and subtractions only!),

AF algorithm routine

- The algorithm seeks the **maximum** of the (noised) **image variance**.
- The focus functions is **global!**
- Various (bi-)**orthogonal expansions** can be used to estimate the variance:
  - trigonometric (DCT, *Fourier*, *Hartley*),
  - *Walsh-Hadamard* (additions and subtractions only!),
  - wavelet, *e.g.* orthogonal *Haar*, *Daubechies*, or biorthogonal *LeGall (5/3)* and *Cohen-Daubechies-Vial (9/7)*,

AF algorithm routine

- The algorithm seeks the **maximum** of the (noised) **image variance**.
- The focus functions is **global!**
- Various (bi-)**orthogonal expansions** can be used to estimate the variance:
  - trigonometric (DCT, *Fourier*, *Hartley*),
  - *Walsh-Hadamard* (additions and subtractions only!),
  - wavelet, *e.g.* orthogonal *Haar*, *Daubechies*, or biorthogonal *LeGall (5/3)* and *Cohen-Daubechies-Vial (9/7)*,
  - polynomial, *e.g.* *Chebyshev*, *Legendre*, *Zernike* (in general—any *'people's polynomials'*).

Possible AF function computation implementations

The following discrete orthogonal series transforms are available in the transform coders:

- *DCT* transform, (**JPEG**),

Possible AF function computation implementations

The following discrete orthogonal series transforms are available in the transform coders:

- *DCT* transform, (**JPEG**),
- *Haar wavelet* transform (**JPEG 2K (Part II)**), and

Possible AF function computation implementations

The following discrete orthogonal series transforms are available in the transform coders:

- *DCT* transform, (**JPEG**),
- *Haar wavelet* transform (**JPEG 2K (Part II)**), and
- *Walsh-Hadamard* transform (**JPEG XR**).

Maximum search algorithms

We can apply standard algorithms to find the **function's maximum** in a **noisy environment**:

- *Golden section-search* (**GSS**) performed on:

Maximum search algorithms

We can apply standard algorithms to find the **function's maximum** in a **noisy environment**:

- *Golden section-search* (**GSS**) performed on:
  - averaged image, or

Maximum search algorithms

We can apply standard algorithms to find the **function's maximum** in a **noisy environment**:

- *Golden section-search* (**GSS**) performed on:
  - averaged image, or
  - smoothed image (*e.g.* by any *de-noising* routine).

Maximum search algorithms

We can apply standard algorithms to find the **function's maximum** in a **noisy environment**:

- *Golden section-search* (**GSS**) performed on:
  - averaged image, or
  - smoothed image (*e.g.* by any *de-noising* routine).
- *Stochastic approximation* (**SA**) exploiting:

Maximum search algorithms

We can apply standard algorithms to find the **function's maximum** in a **noisy environment**:

- *Golden section-search* (**GSS**) performed on:
  - averaged image, or
  - smoothed image (*e.g.* by any *de-noising* routine).
- *Stochastic approximation* (**SA**) exploiting:
  - smoothed image, or

Maximum search algorithms

We can apply standard algorithms to find the **function's maximum** in a **noisy environment**:

- *Golden section-search* (**GSS**) performed on:
  - averaged image, or
  - smoothed image (*e.g.* by any *de-noising* routine).
- *Stochastic approximation* (**SA**) exploiting:
  - smoothed image, or
  - smoothed focus function (*e.g.* by using standard *kernel convolutions*).

AF criteria

1. **Unimodality** – holds in theory. In practice, unimodality can be lost (aperture control!).

AF criteria

1. **Unimodality** – holds in theory. In practice, unimodality can be lost (aperture control!).
2. **Accuracy** – corresponds to the resolution of the sensor.

AF criteria

1. **Unimodality** – holds in theory. In practice, unimodality can be lost (aperture control!).
2. **Accuracy** – corresponds to the resolution of the sensor.
3. **Reproducibility** – a sharp top of the extremum holds in theory.

AF criteria

1. **Unimodality** – holds in theory. In practice, unimodality can be lost (aperture control!).
2. **Accuracy** – corresponds to the resolution of the sensor.
3. **Reproducibility** – a sharp top of the extremum holds in theory.
4. **Range** – global. The variance of the image does not vanish.

AF criteria

1. **Unimodality** – holds in theory. In practice, unimodality can be lost (aperture control!).
2. **Accuracy** – corresponds to the resolution of the sensor.
3. **Reproducibility** – a sharp top of the extremum holds in theory.
4. **Range** – global. The variance of the image does not vanish.
5. **General applicability** – a generic class of processes is admitted (ARMA models, *Markov* fields, and piecewise-smooth models).

AF criteria

1. **Unimodality** – holds in theory. In practice, unimodality can be lost (aperture control!).
2. **Accuracy** – corresponds to the resolution of the sensor.
3. **Reproducibility** – a sharp top of the extremum holds in theory.
4. **Range** – global. The variance of the image does not vanish.
5. **General applicability** – a generic class of processes is admitted (ARMA models, *Markov* fields, and piecewise-smooth models).
6. **Insensitivity to other parameters** – particularly robust against the noise.

AF criteria

1. **Unimodality** – holds in theory. In practice, unimodality can be lost (aperture control!).
2. **Accuracy** – corresponds to the resolution of the sensor.
3. **Reproducibility** – a sharp top of the extremum holds in theory.
4. **Range** – global. The variance of the image does not vanish.
5. **General applicability** – a generic class of processes is admitted (ARMA models, *Markov* fields, and piecewise-smooth models).
6. **Insensitivity to other parameters** – particularly robust against the noise.
7. **Video signal compatibility** – holds

AF criteria

1. **Unimodality** – holds in theory. In practice, unimodality can be lost (aperture control!).
2. **Accuracy** – corresponds to the resolution of the sensor.
3. **Reproducibility** – a sharp top of the extremum holds in theory.
4. **Range** – global. The variance of the image does not vanish.
5. **General applicability** – a generic class of processes is admitted (ARMA models, *Markov* fields, and piecewise-smooth models).
6. **Insensitivity to other parameters** – particularly robust against the noise.
7. **Video signal compatibility** – holds
   - *Block samplers* = *Foveon X3* sensor.

AF criteria

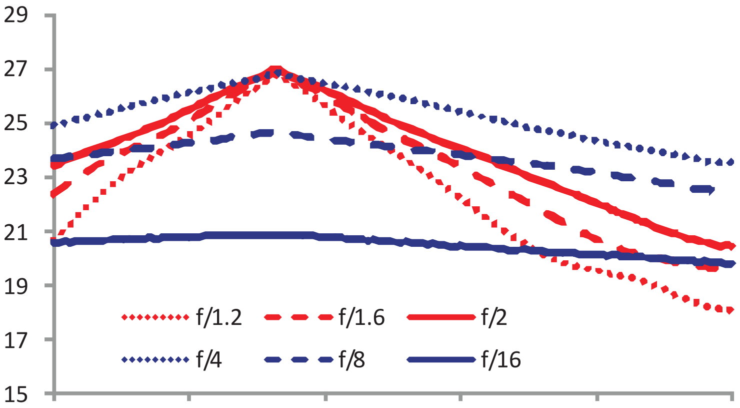
1. **Unimodality** – holds in theory. In practice, unimodality can be lost (aperture control!).
2. **Accuracy** – corresponds to the resolution of the sensor.
3. **Reproducibility** – a sharp top of the extremum holds in theory.
4. **Range** – global. The variance of the image does not vanish.
5. **General applicability** – a generic class of processes is admitted (ARMA models, *Markov* fields, and piecewise-smooth models).
6. **Insensitivity to other parameters** – particularly robust against the noise.
7. **Video signal compatibility** – holds
   - *Block samplers* = *Foveon X3* sensor.
   - *Low-pass filter* + *Dirac comb*-based sampler = *Bayer* CFA sensors.

AF criteria

1. **Unimodality** – holds in theory. In practice, unimodality can be lost (aperture control!).
2. **Accuracy** – corresponds to the resolution of the sensor.
3. **Reproducibility** – a sharp top of the extremum holds in theory.
4. **Range** – global. The variance of the image does not vanish.
5. **General applicability** – a generic class of processes is admitted (ARMA models, *Markov* fields, and piecewise-smooth models).
6. **Insensitivity to other parameters** – particularly robust against the noise.
7. **Video signal compatibility** – holds
   - *Block samplers* = *Foveon X3* sensor.
   - *Low-pass filter* + *Dirac comb*-based sampler = *Bayer* CFA sensors.
8. **Fast implementation** – all algorithms exploit '*fast*' transforms

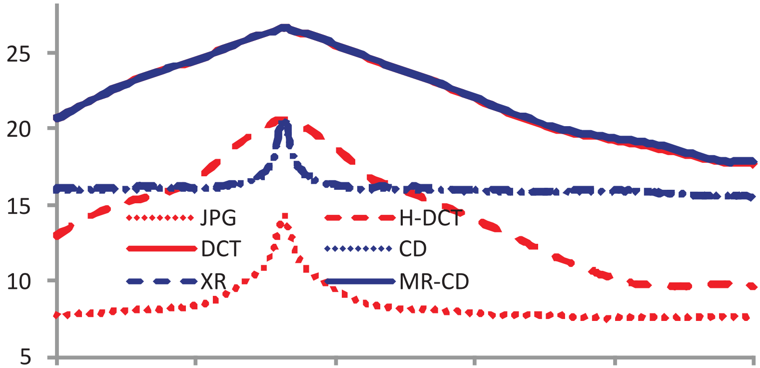
Experimental results

AF against aperture



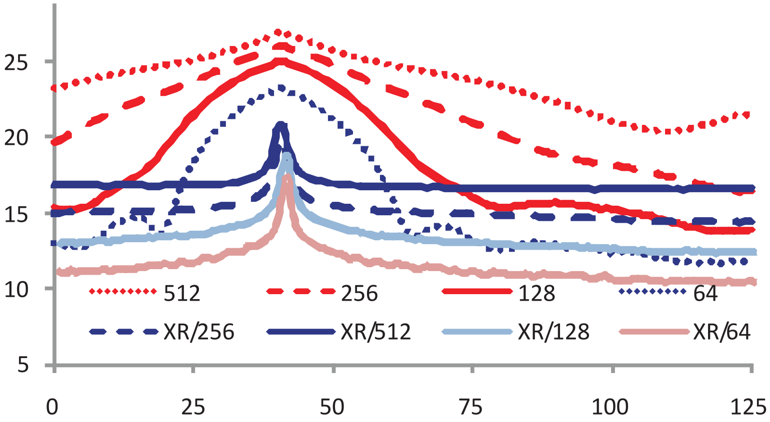
Experimental results

AF against transform coder



Experimental results

AF against image size



Single-Photon AF

Problem

Can the generic algorithm be adapted to the Single-Photon Imagery?

- There are several noise sources

$$\begin{aligned} Y_{lk} &= I_{lk} + Poisson_{lk} + Gaussian_{lk} + PRNU_{lk} \\ &\quad + crosstalk_{lk} + quantization_{lk} + \ldots \end{aligned}$$

Single-Photon AF

Problem

Can the generic algorithm be adapted to the Single-Photon Imagery?

- There are several noise sources

$$\begin{aligned} Y_{lk} &= I_{lk} + Poisson_{lk} + Gaussian_{lk} + PRNU_{lk} \\ &\quad + crosstalk_{lk} + quantization_{lk} + \ldots \end{aligned}$$

- Some of them are **pure** (*random*) noises, *e.g.*:

Single-Photon AF

Problem

Can the generic algorithm be adapted to the Single-Photon Imagery?

- There are several noise sources

$$\begin{aligned} Y_{lk} &= I_{lk} + Poisson_{lk} + Gaussian_{lk} + PRNU_{lk} \\ &\quad + crosstalk_{lk} + quantization_{lk} + \ldots \end{aligned}$$

- Some of them are **pure** (*random*) noises, *e.g.*:
  - Shot noise (of *Poisson* distribution),

Single-Photon AF

Problem

Can the generic algorithm be adapted to the Single-Photon Imagery?

- There are several noise sources

$$
\begin{aligned}
Y_{lk} &= I_{lk} + Poisson_{lk} + Gaussian_{lk} + PRNU_{lk} \\
&\quad + crosstalk_{lk} + quantization_{lk} + \dots
\end{aligned}
$$

- Some of them are **pure** (*random*) noises, *e.g.*:
    - Shot noise (of *Poisson* distribution),
    - Thermal noise (of *Gaussian* distribution).

Single-Photon AF

Problem

Can the generic algorithm be adapted to the Single-Photon Imagery?

- There are several noise sources

$$\begin{aligned} Y_{lk} &= I_{lk} + Poisson_{lk} + Gaussian_{lk} + PRNU_{lk} \\ &\quad + crosstalk_{lk} + quantization_{lk} + \dots \end{aligned}$$

- Some of them are **pure** (*random*) noises, *e.g.*:
    - Shot noise (of *Poisson* distribution),
    - Thermal noise (of *Gaussian* distribution).
- Some are random but fixed, *e.g.* *Photo-Response Non-uniformity*

Single-Photon AF

Problem

Can the generic algorithm be adapted to the Single-Photon Imagery?

- There are several noise sources

$$\begin{aligned} Y_{lk} &= I_{lk} + Poisson_{lk} + Gaussian_{lk} + PRNU_{lk} \\ &\quad + crosstalk_{lk} + quantization_{lk} + \dots \end{aligned}$$

- Some of them are **pure** (*random*) noises, *e.g.*:
  - Shot noise (of *Poisson* distribution),
  - Thermal noise (of *Gaussian* distribution).
- Some are random but fixed, *e.g.* *Photo-Response Non-uniformity*
- For the others, it is just convenient to model them as a noise. . .

Unbalanced Haar Transform

- In order to model **PRNU** for each pixel we use the *Unbalanced Haar Transform* instead of the classic one

Unbalanced Haar Transform

- In order to model **PRNU** for each pixel we use the *Unbalanced Haar Transform* instead of the classic one
- The basic transform step becomes a little bit more complicated than

$$\hat{\alpha}_{m-1,n} = \frac{\sqrt{2}}{2}\hat{\alpha}_{m,2n} + \frac{\sqrt{2}}{2}\hat{\alpha}_{m,2n+1}$$

vs.

$$\bar{\alpha}_{m-1,n} = \sqrt{\frac{I_{m,2n}}{I_{m-1,n}}}\bar{\alpha}_{m,2n} + \sqrt{\frac{I_{m,2n+1}}{I_{m-1,n}}}\bar{\alpha}_{m,2n+1},$$

with $I_{m-1,n} = I_{m,2n+1} + I_{m,2n+1}$ (where $I_{m,2n+1}, I_{m-1,n}$ are the *non-uniformity indices*).

Unbalanced Haar Transform

Application of **UHT** has some advantages**:**

- Can be *plugged-in* into the standard AF algorithm.

Unbalanced Haar Transform

Application of **UHT** has some advantages**:**

- Can be *plugged-in* into the standard AF algorithm.
- Remains fast, *i.e.* linear with number of pixels.

Unbalanced Haar Transform

Application of **UHT** has some advantages**:**

- Can be *plugged-in* into the standard AF algorithm.
- Remains fast, *i.e.* linear with number of pixels.
- Allows for *in situ* image denoising.

Unbalanced Haar Transform

Application of **UHT** has some advantages**:**

- Can be *plugged-in* into the standard AF algorithm.
- Remains fast, *i.e.* linear with number of pixels.
- Allows for *in situ* image denoising.
- Can be computed in parallel.

Unbalanced Haar Transform

Application of **UHT** has some advantages**:**

- Can be *plugged-in* into the standard AF algorithm.
- Remains fast, *i.e.* linear with number of pixels.
- Allows for *in situ* image denoising.
- Can be computed in parallel.
- But. . . requires computing **square roots**. . .

Single-photon AF

We have now

$$Y_{lk} \sim I_{lk} + Poisson_{lk} + Gaussian_{lk}.$$

- The *single-photon-denoising* algorithm has two simple steps:

Single-photon AF

We have now

$$Y_{lk} \sim I_{lk} + Poisson_{lk} + Gaussian_{lk}.$$

- The *single-photon-denoising* algorithm has two simple steps:
  - 'removal' of the Gaussian part by UHT transform with a thresholding. Then

$$Y_{lk} \sim I_{lk} + Poisson_{lk}$$

Single-photon AF

We have now

$$Y_{lk} \sim I_{lk} + Poisson_{lk} + Gaussian_{lk}.$$

- The *single-photon-denoising* algorithm has two simple steps:
  - 'removal' of the Gaussian part by UHT transform with a thresholding. Then

    $$Y_{lk} \sim I_{lk} + Poisson_{lk}$$

  - application of the *Anscombe* transform and repeating the previous step (*i.e.* the UHT transform with a thresholding). Thus

    $$Y_{lk} \sim I_{lk}$$

Single-photon AF

We have now

$$Y_{lk} \sim I_{lk} + Poisson_{lk} + Gaussian_{lk}.$$

- The *single-photon-denoising* algorithm has two simple steps:
  - 'removal' of the Gaussian part by UHT transform with a thresholding. Then

    $$Y_{lk} \sim I_{lk} + Poisson_{lk}$$

  - application of the *Anscombe* transform and repeating the previous step (*i.e.* the UHT transform with a thresholding). Thus

    $$Y_{lk} \sim I_{lk}$$

- The rest of the **AF** algorithm remains unchanged.

Final conclusions

The proposed **AF algorithm:**

- Is *robust* against a noise.

Final conclusions

The proposed **AF algorithm:**

- Is *robust* against a noise.
- Works with a standard (cheap) equipment.

Final conclusions

The proposed **AF algorithm:**

- Is *robust* against a noise.
- Works with a standard (cheap) equipment.
- Can *reuse* existing IPs.

Final conclusions

The proposed **AF algorithm:**

- Is *robust* against a noise.
- Works with a standard (cheap) equipment.
- Can *reuse* existing IPs.
- Can *effectively* be implemented (*e.g. in situ*).

Example

The demonstration movie can be found at:

`http://diuna.iiar.pwr.wroc.pl/sliwinski/gss-af.avi`